

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:27:18 ; Search time 169,442 Seconds  
(without alignments)  
8602.339 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 90  
Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAAGAAAGA 90

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estnu:\*  
4: em\_estov:\*  
5: em\_estpl:\*  
6: em\_estro:\*  
7: em\_estro:\*  
8: em\_estro:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	90	100.0	562	12	BG495857 602540248
2	90	100.0	711	12	BG506209 601860153
3	90	100.0	724	14	BM721352 U1-E-E01-
4	90	100.0	772	12	BG034660 602300022
5	90	100.0	869	12	BG612651 602640078
6	90	100.0	1014	14	BQ221144 AGENCOURT

7	90	100.0	1088	14	BM800217 AGENCOURT
8	88.4	98.2	674	10	AV685706
9	86	95.6	579	14	W27852 4143 Human
10	84.4	93.8	987	9	AL530574
11	78.8	87.6	485	10	BE477083
12	78.6	87.3	922	9	AL530575
13	78	86.7	724	12	BF207826
14	74	82.2	468	10	BE332036
15	74	82.2	557	12	BE851413
16	74	82.2	762	13	BI685252 603310265
17	74	82.2	766	13	BI107517
18	74	82.2	780	14	BQ571052 U1-M-FB0-
19	74	82.2	2061	11	AK003556 Mus muscu
20	73	81.1	785	12	BF133971 601778414
21	72.4	80.7	1951	11	AK010176 Mus muscu
22	61.8	68.4	242	9	AV218623 AV218623
23	51.4	57.1	336	10	AV695603 AV695603
24	48.2	53.6	836	12	BF185342 601844161
25	46	51.1	438	9	AL698051 DREF2P686C
26	43.2	48.0	632	13	BJ026001 BJ026001
27	42.4	47.1	718	13	BJ016230 BJ016230
28	38.4	42.7	777	13	BM798211 K-EST0081
29	36.9	44.2	442	10	BE531754 u055C10.Y
30	33.2	35.3	479	17	A2054089 RPT1-23-4
31	31.8	34.4	473	17	GC0000548 Gallus ga
32	31	34.2	513	13	BI808143 C003D09 O
33	30.8	34.2	700	12	BG827372 602749288
34	30.2	33.6	1101	17	CNS00KX0 Drosophi1
35	29.8	32.7	666	10	BE261674 601194436
36	29.4	32.1	1052	17	CNS010HY Drosophi1
37	29	32.2	354	13	BI062622 IL3-UT011
38	28.4	31.6	503	12	BG168700 602319956
39	28.2	31.3	758	13	BI668281 603392238
40	28.2	31.3	396	10	AM024086 wu62c06.x
41	28.2	31.3	439	14	BU024120 OHF1CF19.
42	28.2	31.3	512	10	AM509578 ga61e12.Y
43	28.2	31.3	525	9	AU241057 AU241057
44	28.2	31.3	590	13	BJ007993 BJ007993
45	28.2	31.3	590	13	BJ007993 BJ007993

#### ALIGNMENTS

RESULT 1  
BG495857  
LOCUS 562 bp mRNA linear EST 27-MAR-2001  
DEFINITION 602540248F1 NIH\_MGC\_59 Homo sapiens cDNA clone IMAGE:4671566 5',  
mRNA sequence.  
ACCESSION BG495857  
VERSION BG495857.1 GI:13457373  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 562)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM486 row: 0 column: 11  
High quality sequence stop: 562.  
Location/Qualifiers  
1..562

#### FEATURES

source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4671586"
/clone_lib="NIH_MGC_59"
/tissue_type="mucoepidermoid carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattagcc);
Double stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATAGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGAGCGCCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT      167 a      122 c      142 g      131 t
ORIGIN
Query Match      100.0%; Score 90; DB 12; Length 562;
Best Local Similarity 100.0%; Pred. No. 8.9e-20;
Matches 90; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

OY 1 TGTGACCTGTGATCGAATCATCTATGGGATCGCGACGCGACATAAATCC 60
|||||
DB 205 TGTGACCTGTGATCGAATCATCTATGGGATCGCGACGCGACATAAATCC 264
|||||

OY 61 AATCCCACTTGAACTGACACTGAGAAAGA 90
|||||
DB 265 AATCCCACTTGAACTGACACTGAGAAAGA 294
|||||

RESULT 2
LOCUS      BG506209      711 bp      mRNA      linear      EST 27-MAR-2001
DEFINITION 601860153P1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4072428 5',
mRNA sequence.
ACCESSION  BG506209
VERSION    BG506209.1  GI:13467726
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 711)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS   Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
JOURNAL   Tissue Procurement: ATCC
COMMENT   cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            DNA Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1CM919 row: j column: 13
            High quality sequence stop: 643.
            Location/Qualifiers
                1..711
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="IMAGE:4072428"
                    /clone_lib="NIH_MGC_61"
                    /tissue_type="embryonal carcinoma"
                    /lab_host="DH10B (TI phage-resistant)"
                    /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1:
                    SfiI (ggcgccctcgcc); Site_2: SfiI (ggccattagcc);
                    Double stranded cDNA was prepared from cell line RNA. 5'
                    and 3' adaptors were used in cloning as follows: 5'
```

```
adaptor sequence: 5'-CACGGCCATTATAGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGAGCGCGCGACATG-dt(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT      224 a      145 c      165 g      177 t
ORIGIN
Query Match      100.0%; Score 90; DB 12; Length 711;
Best Local Similarity 100.0%; Pred. No. 9.4e-20;
Matches 90; Conservative 0; Mismatches 0; Gaps 0;
Indels 0;

OY 1 TGTGACCTGTGATCGAATCATCTATGGGATCGCGACGCGACATAAATCC 60
|||||
DB 74 TGTGACCTGTGATCGAATCATCTATGGGATCGCGACGCGACATAAATCC 133
|||||

OY 61 AATCCCACTTGAACTGACACTGAGAAAGA 90
|||||
DB 134 AATCCCACTTGAACTGACACTGAGAAAGA 163
|||||

RESULT 3
LOCUS      BM721352      724 bp      mRNA      linear      EST 01-MAR-2002
DEFINITION UI-E-E01-1b-b-20-0-UI-r1 UI-E-E01 Homo sapiens cDNA clone
UI-E-E01-1b-b-20-0-UI 5', mRNA sequence.
ACCESSION  BM721352
VERSION    BM721352.1  GI:19041207
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE  1 (bases 1 to 724)
            Bonaldo,M.F., Lennon,G. and Soares,M.B.
            Normalization and subtraction: two approaches to facilitate gene
            discovery
            Genome Res. 6 (9), 791-806 (1996)
JOURNAL   Contact: Soares, MB
            MEDLINE 97044477
COMMENT   Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: mssoares@blue.weeg.iowa.edu
            Tissue Procurement: Dr. Gregg Hageman
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            DNA Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..724
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="UI-E-E01-1b-b-20-0-UI"
                    /clone_lib="UI-E-E01"
                    /tissue_type="fetal eye"
                    /dev_stage="fetal"
                    /lab_host="DH10B (Life Technologies) (TI phage resistant)"
                    /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
                    modified polylinker; Site_1: EcoR I; Site_2: Not I;
                    UI-E-E01 is a normalized cDNA library containing the
                    following tissue(s): fetal eye. The library was
                    constructed according to Bonaldo, Lennon and Soares,
                    Genome Research, 6:791-806, 1996. First strand cDNA
                    synthesis was primed with an oligo-dT primer containing a
                    Not I site. Double stranded cDNA was ligated to an EcoR I
```



adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGCATACC. This library was created for the program Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 208 a 132 c 175 g 208 t 1 others

ORIGIN

Query Match 100.0%; Score 90; DB 14; Length 724;  
Best Local Similarity 100.0%; Pred. No. 9.4e-20;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 60  
|||||  
Db 5 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 64  
|||||

OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90  
|||||  
Db 65 AAATCCCACTTGAACCACTGAAGAAAAGA 94  
|||||

RESULT 4  
BG034660 772 bp mRNA linear EST 24-JAN-2001  
LOCUS 602300022F1 NIH\_MGC\_87 Homo sapiens cDNA clone IMAGE:4394260 5',  
DEFINITION mRNA sequence.  
ACCESSION BG034660  
VERSION BG034660.1 GI:12428227  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 772)  
NIH-MGC http://mgc.ncl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHAM10089 row: 1 column: 05  
High quality sequence stop: 692.

FEATURES  
source Location/Qualifiers  
1. 772  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4394260"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site\_1: NCI; Site\_2: Sali; Cloned unidirectionally; Oligo-dT primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT 224 a 155 c 178 g 214 t 1 others

ORIGIN

Query Match 100.0%; Score 90; DB 12; Length 772;  
Best Local Similarity 100.0%; Pred. No. 9.5e-20;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 60  
|||||  
Db 181 AAATCCCACTTGAACCACTGAAGAAAAGA 210  
|||||

Db 121 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 180  
OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90  
|||||  
Db 181 AAATCCCACTTGAACCACTGAAGAAAAGA 210  
|||||

RESULT 5  
BG612651 869 bp mRNA linear EST 18-APR-2001  
LOCUS 602640078F1 NIH\_MGC\_61 Homo sapiens cDNA clone IMAGE:4771338 5',  
DEFINITION mRNA sequence.  
ACCESSION BG612651  
VERSION BG612651.1 GI:13664022  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 869)  
NIH-MGC http://mgc.ncl.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: c9apbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LHCMI639 row: k column: 19  
High quality sequence stop: 684.

FEATURES  
source Location/Qualifiers  
1. 869  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4771338"  
/tissue\_type="embryonal carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site\_1: SfiI (99ccgctcgagc); Site\_2: SfiI (99ccatattgccc); SfiI (99ccgctcgagc); Site\_2: SfiI (99ccatattgccc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGCAGCGCGCCGACATG-dT(30)-BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

BASE COUNT 259 a 184 c 206 g 220 t

ORIGIN

Query Match 100.0%; Score 90; DB 12; Length 869;  
Best Local Similarity 100.0%; Pred. No. 9.8e-20;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 60  
|||||  
Db 74 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGGCACATATAATCC 133  
|||||

OY 61 AAATCCCACTTGAACCACTGAAGAAAAGA 90  
|||||  
Db 134 AAATCCCACTTGAACCACTGAAGAAAAGA 163  
|||||

RESULT 6  
BQ221144 1014 bp mRNA linear EST 02-MAY-2002  
LOCUS BQ221144

DEFINITION AGENCOURT\_7590416 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6065448  
5', mRNA sequence.  
ACCESSION BQ221144  
VERSION BQ221144.1 GI:20402544  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 1014)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM13341 row: m column: 01  
High quality sequence stop: 387.

FEATURES  
source  
1..1014  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6065448"  
/clone\_1lb="NIH\_MGC\_72"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: skin; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

BASE COUNT 289 a 283 c 231 g 211 t  
ORIGIN

Query Match 100.0%; Score 90; DB 14; Length 1014;  
Best Local Similarity 100.0%; Pred. No. 1e-19;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCATCATCTGGGATGGGACGACGACATTAATCC 60  
|||||  
Db 295 TGTGACCTCTGTGATCATCATCTGGGATGGGACGACGACATTAATCC 354  
|||||  
OY 61 AATCCACTTGAACCACTGAAGAAGA 90  
|||||  
Db 355 AATCCACTTGAACCACTGAAGAAGA 384  
|||||

RESULT 7  
LOCUS BM800217 1088 bp mRNA linear EST 05-MAR-2002  
DEFINITION AGENCOURT\_6416157 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5531248  
5', mRNA sequence.  
ACCESSION BM800217  
VERSION BM800217.1 GI:191117040  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 1088)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM12212 row: j column: 17  
High quality sequence stop: 651.

FEATURES  
source  
1..1088  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5531248"  
/clone\_1lb="NIH\_MGC\_71"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."  
High quality sequence stop: 651.

BASE COUNT 318 a 218 c 255 g 294 t 3 others  
ORIGIN

Query Match 100.0%; Score 90; DB 14; Length 1088;  
Best Local Similarity 100.0%; Pred. No. 1e-19;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCATCATCTGGGATGGGACGACGACATTAATCC 60  
|||||  
Db 283 TGTGACCTCTGTGATCATCATCTGGGATGGGACGACGACATTAATCC 342  
|||||  
OY 61 AATCCACTTGAACCACTGAAGAAGA 90  
|||||  
Db 343 AATCCACTTGAACCACTGAAGAAGA 372  
|||||

RESULT 8  
LOCUS AV685706 674 bp mRNA linear EST 16-JAN-2002  
DEFINITION AV685706 GKC Homo sapiens cDNA clone GKCHYG04 5', mRNA sequence.  
ACCESSION AV685706  
VERSION AV685706.1 GI:10287569  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 674)  
AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,  
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Li, N., Du, J., Hu, W.,  
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,  
Hu, G., Gu, J., Chen, Z., and Han, Z.  
Insight into hepatocellular carcinogenesis at transcriptome level  
by comparing gene expression profiles of hepatocellular carcinoma  
with those of corresponding noncancerous liver  
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL MEDLINE  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..674  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="GKCHYG04"  
/clone\_1lb="GKC"  
/tissue\_type="hepatocellular carcinoma"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

```

BASE COUNT      206 a      141 c      163 g      157 t      7 others
ORIGIN
Query Match      98.2%; Score 88.4; DB 10; Length 674;
Best Local Similarity 98.9%; Pred. No. 3.2e-19;
Matches 89; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATGCAATCATCTTTGGGATCGCGAATGGCAGCGCACATAAAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 323 TGTGACCCCTGTGATGCAATCATCTTTGGGATCGCGAATGGCAGCGCACATAAAATCC 382
OY 61 AAATCCCACTTGAACCACTGAAGAAAGA 90
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 383 AAATCCCACTTGAACCACTGAAGAAAGA 412

RESULT 9
W27852 579 bp mRNA linear EST 08-MAY-1996
LOCUS W27852.1 GI:1308007
DEFINITION 41d3 Human retina cDNA randomly primed sublibrary Homo sapiens cDNA
          mRNA sequence.
ACCESSION W27852
VERSION W27852.1 GI:1308007
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 579)
AUTHORS Macke, J., Smallwood, P. and Nathans, J.
TITLE Adult Human Retina cDNA
JOURNAL Unpublished (1996)
COMMENT Contact: Dr. Jeremy Nathans
          Dr. Jeremy Nathans, Dept. of Molecular Biology and Genetics
          Johns Hopkins School of Medicine
          725 North Wolfe Street, Baltimore, MD 21205
          Tel: 410 955 4678
          Fax: 410 614 0827
          Email: jeremy.nathans@jhu.edu
          Clones from this library are NOT available.
          PCR Primers
          FORWARD: CTTTGGAGCAAGTTCAAGCCCTGTTAACT
          BACKWARD: GAGTGCTTATGATGATTTCTCCAGCGTAA
          Seq primer: GGGTAAACCAAGCAAGATT.

FEATURES
    source
        location/Qualifiers
            1..579
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_idb="Human retina cDNA randomly primed sublibrary"
                /sex="mixed (males and females)"
                /tissue_type="retina"
                /dev_stage="adult"
                /lab_host="E. coli strain K802"
                /note="Organ: eye; Vector: lambda gt10; Site_1: EcoRI;
                Site_2: EcoRI; The library used for sequencing was a
                sublibrary derived from a human retina cDNA library.
                Inserts from retina cDNA library DNA were isolated.
                Randomly primed, PCR amplified, size-selected, and cloned
                into lambda gt10. Individual plaques were arrayed and used
                as templates for PCR amplification, and these PCR
                products were used for sequencing."
BASE COUNT      97 a      110 c      91 g      136 t      145 others
ORIGIN
Query Match      95.6%; Score 86; DB 14; Length 579;
Best Local Similarity 96.6%; Pred. No. 2e-18;
Matches 86; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATGCAATCATCTTTGGGATCGCGAATGGCAGCGCACATAAAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 318 TGTGACCTCTGTGATGCAATCATCTTTGGGATCGCGAATGGCAGCGCACATAAAATCC 259
OY 61 AAATCCCACTTGAACCACTGAAGAAAG 89

```

```

DB 258 AAATCCCACTTGAACCACTGAGAAAG 230
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 10
LOCUS AL530574/c 987 bp mRNA linear EST 13-FEB-2001
DEFINITION AL530574 LTI_NFL001_NBC4 Homo sapiens cDNA clone CSDD0081N11 3
          prime, mRNA sequence.
ACCESSION AL530574
VERSION AL530574.1 GI:12794067
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 987)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    source
        location/Qualifiers
            1..987
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_idb="CSDD0081N11"
                /clone_idb="LTI_NFL001_NBC4"
                /sex="male"
                /tissue_type="neuroblastoma cells"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                was primed with a NotI-oligo(dT) primer. Five prime end
                enriched, double-stranded cDNA was digested with Not I and
                cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                vector. Library was normalized. Library was constructed
                by Life Technologies, a division of Invitrogen 9800 Medical Center
                Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                8371 Email : filangel@fatech.com URL :
                http://fulllength.invitrogen.com"
BASE COUNT      281 a      230 c      180 g      284 t      12 others
ORIGIN
Query Match      93.8%; Score 84.4; DB 9; Length 987;
Best Local Similarity 94.4%; Pred. No. 7.6e-18;
Matches 85; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATGCAATCATCTTTGGGATCGCGAATGGCAGCGCACATAAAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 885 TGTGACCTCTGTGATGCAATCATCTTTGGGATCGCGAATGGCAGCGCACATAAAATCC 826
OY 61 AAATCCCACTTGAACCACTGAAGAAAGA 90
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 825 AAATCCCACTTGAACCACTGAAGAAAGR 796

RESULT 11
LOCUS BE477083 485 bp mRNA linear EST 28-AUG-2000
DEFINITION BE477083 160729 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE477083
VERSION BE477083.1 GI:9596688
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 485)
          Sonstegard, T.S., Capuco, A.V., Van Tassel, C.P., Ashwell, M.S. and

```

TITLE Wells, R.D.  
 JOURNAL Mapping of Expressed Sequence Tags from a normalized bovine mammary gland cDNA library  
 COMMENT Unpublished (2000)  
 Contact: Sonstegard TS  
 USDA, ARS, Beltsville Agricultural Research Center  
 Bldg. 200, Rm 2A, Beltsville, MD 20705, USA  
 Tel: 301 504 8416  
 Fax: 301 504 8414  
 Email: tads@arri.barc.usda.gov  
 Single pass sequencing. Bases called and alt\_trimmed with phred v0.980904.e. Vector identified by cross\_match with the -minscore 18 and -mismatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCATGACCAT  
 BACKWARD: CTTTCCCATGACGACG  
 Plate: 144 Row: D Column: 7  
 Seq primer: ATTAGTGACCATATAG.  
 Location/Qualifiers  
 1..485  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="BARC 5BOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."  
 BASE COUNT 141 a 97 c 120 g 127 t  
 ORIGIN  
 Query Match 87.6%; Score 78.8; DB 10; Length 485;  
 Best Local Similarity 92.2%; Pred. No. 4.9e-16;  
 Matches 83; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 Oy 1 TGTGACCTCTGTGATGCAATCATCATTTGGGATGGCGACGCCACATATAATCC 60  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 87 TGTGACCTCTGTGATGCAATCATCATTTGGGATGGCGACGCCACATATAATCC 146  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Oy 61 AATGCCACTTGACCACTGAGAAAGA 90  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 147 AATGCCACTTGACCACTGAGAAAGA 176  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 RESULT 12  
 LOCUS AL530575 922 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL530575 LTI\_NFL001\_NBC4 Homo sapiens cDNA clone CS0DD008YNI1 5 prime, mRNA sequence.  
 ACCESSION AL530575  
 VERSION AL530575.1 GI:12794068  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 922)  
 AUTHORS Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.  
 Location/Qualifiers  
 1..922  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DD008YNI1"  
 /clone\_lib="LTI\_NFL001\_NBC4"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"

/lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : filiang@lifestech.com URL : http://fulllength.invitrogen.com"  
 BASE COUNT 258 a 210 c 235 g 215 t 4 others  
 ORIGIN  
 Query Match 87.3%; Score 78.6; DB 9; Length 922;  
 Best Local Similarity 97.8%; Pred. No. 6.6e-16;  
 Matches 89; Conservative 1; Mismatches 0; Indels 1; Gaps 1;  
 Oy 1 TGTGACCTCTGTGATGCAATCATCATTTGGGATGGCGACGCCACATATAATC 59  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 793 TGTGACCTCTGTGATGCAATCATCATTTGGGATGGCGACGCCACATATAATC 852  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Oy 60 CAATGCCACTTGACCACTGAGAAAGA 90  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 853 CAATGCCACTTGACCACTGAGAAAGA 883  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 RESULT 13  
 LOCUS BF207826 724 bp mRNA linear EST 06-NOV-2000  
 DEFINITION BF207826 601862530F1 NIH\_MGC\_53 Homo sapiens cDNA clone IMAGE:4082070 5', mRNA sequence.  
 ACCESSION BF207826  
 VERSION BF207826.1 GI:11101412  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 724)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs.rem.nhl.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Clontech Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov  
 Plate: LNCM944 row: 1 column: 07  
 High quality sequence stop: 631.  
 Location/Qualifiers  
 1..724  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4082070"  
 /clone\_lib="NIH\_MGC\_53"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: bladder; Vector: pDRR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcgctatggcc)"; Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCACTTATGGCC-3' and 3' adaptor sequence: 5'-ATCTAGAGCCGACGCCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT

208 a 153 c 177 g 186 t

```

ORIGIN
Query Match      86.7%; Score 78; DB 12; Length 724;
Best Local Similarity 98.9%; Pred. No. 9.9e-16;
Matches 89; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGCAATGGCAGCCGACATATAATCC 60
    |||||||
DB 202 TGTGACCTCTGTGATCGAATCATCA-TGGGATCGGCAATGGCAGCCGACATATAATCC 260
    |||||||

OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
    |||||||
DB 261 AATCCCACTTGAACCACTGAAGAAAAGA 290
    |||||||

RESULT 14
BE332036      468 bp      mRNA      linear      EST 14-JUL-2000
LOCUS         ut07a11.y1 Soares.thymus.2NDBMT Mus musculus cDNA clone
DEFINITION    IMAGE:3327164.5, similar to WP:2C395.6 CE01442 TRNA
               ISOENTENYLTRANSFERASE ;, mRNA sequence.
ACCESSION     BE332036
VERSION       BE332036.1 GI:9205812
KEYWORDS      EST.
SOURCE        house mouse.
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 468)
AUTHORS       NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               This clone is available royalty-free through LNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:1071328
               Seq primer: -40RP from Gibco
               High quality sequence stop: 462.
               Location/Qualifiers
               1. 468
               /organism="Mus musculus"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="IMAGE:3327164"
               /clone_1lb="Soares.Chymus_2NDBMT"
               /sex="male"
               /tissue_type="Thymus"
               /dev_stage="4 weeks"
               /lab_host="DH10B"
               /note="Vector: pT773D-Pac (Pharmacia) with a modified
               polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
               was primed with a Not I - oligo(dT) primer (5'
               TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGCAATGGCAGCCGACATATAATCC
               3'); double-stranded cDNA was ligated to Eco RI adaptors
               (Pharmacia), digested with Not I and cloned into the Not I
               and Eco RI sites of the modified pT773 vector. RNA
               provided by Dr. Bertrand Jordan. Library went through two
               rounds of normalization, and was constructed by Bento
               Soares and M. Fatima Bonaldo."

BASE COUNT      148 a      101 c      131 g      88 t
ORIGIN

Query Match      82.2%; Score 74; DB 10; Length 468;
Best Local Similarity 88.9%; Pred. No. 2e-14;
Matches 80; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGCAATGGCAGCCGACATATAATCC 60
    |||||||
DB 161 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGCAATGGCAGCCGACATATAATCC 220
    |||||||

OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
    |||||||

```

```

DB 221 AATCTCACTTGCACCAATTAAGAAAAGA 250
    |||||||

RESULT 15
BE851413      557 bp      mRNA      linear      EST 26-SEP-2000
LOCUS         uw95g12.y1 Soares.mammary_gland.NMLMG Mus musculus cDNA clone
DEFINITION    IMAGE:3469798.5, similar to SW:MOD5_YEAST P07884 TRNA
               ISOENTENYLTRANSFERASE ;, mRNA sequence.
ACCESSION     BE851413
VERSION       BE851413.1 GI:10309752
KEYWORDS      EST.
SOURCE        house mouse.
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 557)
AUTHORS       NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE         National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT       Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               This clone is available royalty-free through LNL; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.
               MGI:1389158
               Seq primer: -40RP from Gibco
               High quality sequence stop: 426.
               Location/Qualifiers
               1. 557
               /organism="Mus musculus"
               /db_xref="taxon:10090"
               /clone="IMAGE:3469798"
               /clone_1lb="Soares.mammary_gland.NMLMG"
               /sex="female (lactating)"
               /tissue_type="mammary gland"
               /lab_host="DH10B"
               /note="Vector: pT773D-Pac (Pharmacia) with a modified
               polylinker; 1st strand cDNA was prepared from mammary
               gland tissue from a lactating female, and was then primed
               with a Not I - oligo(dT) primer. Double-stranded cDNA was
               ligated to Eco RI adaptors (Pharmacia), digested with Not
               I and cloned into the Not I and Eco RI sites of the
               modified pT773 vector. Library is normalized. Library
               was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT      173 a      124 c      146 g      114 t
ORIGIN

Query Match      82.2%; Score 74; DB 12; Length 557;
Best Local Similarity 88.9%; Pred. No. 2.1e-14;
Matches 80; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGCAATGGCAGCCGACATATAATCC 60
    |||||||
DB 297 TGTGACCTCTGTGACCGGATCATCATTTGGGATCGGCAATGGCAGCCGACATATAATCC 356
    |||||||

OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
    |||||||
DB 357 AATCTCACTTGCACCAATTAAGAAAAGA 386
    |||||||

```

Search completed: April 21, 2003, 18:21:16  
 Job time : 174.442 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:19:26 ; Search time 5959 Seconds

(Without alignments)  
9967.919 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041

Sequence: 1 CTCGCATTAAGATGCGCGCG.....TTTACAGAGAAAAAAA 2041

Scoring table: OLIGO\_NUC

Gap 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2041	100.0	2041	6	A98668	A98668 Sequence 3
2	1258	61.6	1308	6	AY052768	AY052768 Homo sapi
3	896	43.9	2130	6	AX405704	AX405704 Sequence
4	825	40.4	2216	9	AF074918	AF074918 Homo sapi
5	765	37.5	1844	9	BC010741	BC010741 Homo sapi
6	751	36.8	1749	9	AK000068	AK000068 Homo sapi
7	751	36.8	1749	9	AK074222	AK074222 Homo sapi
8	593	29.1	166496	9	HS118121	AL033522 Homo sapi
9	593	29.1	200484	2	AC025002	AC025002 Homo sapi
10	473	23.2	199863	2	AL354888	AL354888 Homo sapi
11	473	23.2	199863	2	AL354888	AL354888 Homo sapi
12	241	11.8	355	6	AX336538	AX336538 Sequence
13	150	7.3	252	11	G24438	G24438 human STS W
14	143	7.0	199863	2	AL354888	AL354888 Homo sapi
15	120	5.9	79785	2	AC025068	AC025068 Homo sapi
16	48	2.4	2045	10	BC019812	BC019812 Mus muscu
17	48	2.4	198772	10	AL606906	AL606906 Mouse DNA
18	34	1.7	77304	2	AC103185	AC103185 Rattus no
19	34	1.7	201861	2	AC103344	AC103344 Rattus no
20	32	1.6	201861	2	AC103344	AC103344 Rattus no
21	26	1.3	66680	2	AC100483	AC100483 Mus muscu
22	26	1.3	69888	2	AC087576	AC087576 Homo sapi
23	26	1.3	75021	2	AC040889	AC040889 Homo sapi
24	26	1.3	129332	2	AL590028	AL590028 Human DNA
25	26	1.3	145363	2	AC079213	AC079213 Homo sapi
26	26	1.3	149646	2	AL163536	AL163536 Homo sapi
27	26	1.3	160708	2	AC068406	AC068406 Homo sapi
28	26	1.3	161486	2	AL590008	AL590008 Homo sapi
29	26	1.3	163603	2	AC026755	AC026755 Homo sapi
30	26	1.3	195169	2	AC017109	AC017109 Homo sapi
31	26	1.3	197310	9	AC099754	AC099754 Homo sapi
32	26	1.3	198958	2	AC020688	AC020688 Homo sapi
33	25	1.2	250529	9	HUAE000658	AE000658 Homo sapi
34	24	1.2	931767	9	AC007748	AC007748 Homo sapi
35	24	1.2	99593	9	HS0857M17	AL117381 Human DNA
36	24	1.2	157658	2	AC073285	AC073285 Homo sapi
37	24	1.2	158692	9	AC009490	AC009490 Homo sapi
38	24	1.2	160529	9	AC092659	AC092659 Homo sapi
39	24	1.2	162548	2	AC131158	AC131158 Homo sapi
40	24	1.2	163377	2	AC007897	AC007897 Homo sapi
41	24	1.2	166588	2	AC021870	AC021870 Homo sapi
42	24	1.2	176848	2	AP002367	AP002367 Homo sapi
43	24	1.2	176848	9	AC008082	AC008082 Homo sapi
44	24	1.2	175822	2	AC068402	AC068402 Homo sapi
45	24	1.2	192338	2	AC022221	AC022221 Homo sapi
						AC087286 Homo sapi

## ALIGNMENTS

```

RESULT 1
A98668 LOCUS A98668 2041 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 3 from Patent WO9910482.
ACCESSION A98668
VERSION A98668.1 GI:6781710
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 2041)
AUTHORS Lemieux,J. and Hekimi,S.
TITLE THE C. ELEGANS GRO-1 GENE
JOURNAL Patent: WO 9910482-A 3 04-MAR-1999;
LEMIEUX JASON (CA); UNIV MCGILL (CA)

```

FEATURES Location/Qualifiers  
 source 1..2041  
 /organism="unidentified"  
 /db\_xref="taxon:3264"  
 BASE COUNT 589 a 421 c 502 g 529 t  
 ORIGIN

Query Match 100.0%; Score 2041; DB 6; Length 2041;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CTGCCATTAAGATGCGCTGGCGGCTGCACAGAGCATTCCTGTGGCGAGTGGGCTCAG 60
Db 1 CTGCCATTAAGATGCGCTGGCGGCTGCACAGAGCATTCCTGTGGCGAGTGGGCTCAG 60
QY 61 GGGCTGCAAGGAGACCTTACCTCTTGTAGTATCTCGGGCCACGGGCAACGCAATC 120
Db 61 GGGCTGCAAGGAGACCTTACCTCTTGTAGTATCTCGGGCCACGGGCAACGCAATC 120
QY 121 CAGGCTGGGCTTGCAGTACGAGGCGGCTCGGCGGTGAGATGTCAGCGCTGACTCAT 180
Db 121 CAGGCTGGGCTTGCAGTACGAGGCGGCTCGGCGGTGAGATGTCAGCGCTGACTCAT 180
QY 181 GCAGGTCTATGAAGGCTTAGACATCATCACCAAGAGTTTGTGCCAAGAGAGAGAT 240
Db 181 GCAGGTCTATGAAGGCTTAGACATCATCACCAAGAGTTTGTGCCAAGAGAGAGAT 240
QY 241 CTGCGGCGACCATGATGATCAGCTTGTGATCCTCTTGTGACCAATTTACAGTGTGA 300
Db 241 CTGCGGCGACCATGATGATCAGCTTGTGATCCTCTTGTGACCAATTTACAGTGTGA 300
QY 301 CTTCAGAAATAGAGCAACTCTCTGATGAGATATATTTTGGCCGAGACAAATTCCTAT 360
Db 301 CTTCAGAAATAGAGCAACTCTCTGATGAGATATATTTTGGCCGAGACAAATTCCTAT 360
QY 361 TGTGTGGGAGGAGCAACATATTACATGATCTCTGCTGGAAGTTTGTGCAATAC 420
Db 361 TGTGTGGGAGGAGCAACATATTATACATGATCTCTGCTGGAAGTTTGTGCAATAC 420
QY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTGAACCGAAAGTGGACGCTTGAAGAAAGA 480
Db 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTGAACCGAAAGTGGACGCTTGAAGAAAGA 480
QY 481 GGATGGCTGTGATCTTACAAAGCCCTAAGCCAGGTGAGACCAAGAAATGCTGCAAGCT 540
Db 481 GGATGGCTGTGATCTTACAAAGCCCTAAGCCAGGTGAGACCAAGAAATGCTGCAAGCT 540
QY 541 GCATCCACATGACAAAGCGAAAGTGGCCAGAGCTTGCAAGTTTGTGAAGAAACAGAGAT 600
Db 541 GCATCCACATGACAAAGCGAAAGTGGCCAGAGCTTGCAAGTTTGTGAAGAAACAGAGAT 600
QY 601 CTCTCATAGTGAATTTCTCCATCGTCAACATACGGAAGAGTGTGTGCTCCCTTGAGAG 660
Db 601 CTCTCATAGTGAATTTCTCCATCGTCAACATACGGAAGAGTGTGTGCTCCCTTGAGAG 660
QY 661 TCTCTGAAGTCTCTAACCCTTGCACTCTTGGCTTCAGTGCAGCAGGCACTTCTCAGA 720
Db 661 TCTCTGAAGTCTCTAACCCTTGCACTCTTGGCTTCAGTGCAGCAGGCACTTCTCAGA 720
QY 721 TGAGCGCTTGAGATAGAGGCTGATGACATGTTGCTGCTGAGGCTCTTGGAGAACTAG 780
Db 721 TGAGCGCTTGAGATAGAGGCTGATGACATGTTGCTGCTGAGGCTCTTGGAGAACTAG 780
QY 781 AGATTTTACAGAGCGCTATATAGAGAAATGTTTGGAAATATAGCCAGAGCATATCACA 840
Db 781 AGATTTTACAGAGCGCTATATAGAGAAATGTTTGGAAATATAGCCAGAGCATATCACA 840
QY 841 TGATATCTTCATTCATATGCTTCAAGGATTTCAAGGATTTCAAGTACTGTACTCAAGGAAA 900
Db 841 TGATATCTTCATTCATATGCTTCAAGGATTTCAAGGATTTCAAGTACTGTACTCAAGGAAA 900
QY 901 ATGCACACTGGAGACTAGTAAACAGCTTCTTAAGAAAGACCTGTGCTCCATTTGCCCCC 960
  
```

```

Db 901 ATGCACACTGGAGACTAGTAAACAGCTTCTTAAGAAAGACCTGTGCTCCATTTGCCCCC 960
QY 961 TGTCTATGCTTAGAGGTATCTGATGTCTCAGAGTGGAGAGATGCTGTCTTGAACCTGCG 1020
Db 961 TGTCTATGCTTAGAGGTATCTGATGTCTCAGAGTGGAGAGATGCTGTCTTGAACCTGCG 1020
QY 1021 TCTTGAATTCGTCAAAGTTTCAATCCAGGCGCAAGCCCTACAGCCTCAATTAAGAT 1080
Db 1021 TCTTGAATTCGTCAAAGTTTCAATCCAGGCGCAAGCCCTACAGCCTCAATTAAGAT 1080
QY 1081 GCCATACAAATGAAGCTGAGAACAGAAAGTTTATCAGCTGTGACCTCTGTATGCAAT 1140
Db 1081 GCCATACAAATGAAGCTGAGAACAGAAAGTTTATCAGCTGTGACCTCTGTATGCAAT 1140
QY 1141 CATCATTTGGGATCGCGAATGGCGACGCACATTAATATCCAAATCCCACTTGAACCACT 1200
Db 1141 CATCATTTGGGATCGCGAATGGCGACGCACATTAATATCCAAATCCCACTTGAACCACT 1200
QY 1201 GAAGAAAGAGAAAGATTTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTC 1260
Db 1201 GAAGAAAGAGAAAGATTTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTC 1260
QY 1261 CCCAGACTATTAACAAGAACTTAAAGGAGATCCCGAGGCGAGAAATGATCAAGAGCT 1320
Db 1261 CCCAGACTATTAACAAGAACTTAAAGGAGATCCCGAGGCGAGAAATGATCAAGAGCT 1320
QY 1321 GAAATCGAGCTTTAAGAGACATGTCAGTGGCTTTGGAAGAGTGTGGGATCCAGTT 1380
Db 1321 GAAATCGAGCTTTAAGAGACATGTCAGTGGCTTTGGAAGAGTGTGGGATCCAGTT 1380
QY 1381 CAGAGGAGAGGATGTTGTCTCCAGCTGGGCAAGAGATGCTATGCCGAATTTTC 1440
Db 1381 CAGAGGAGAGGATGTTGTCTCCAGCTGGGCAAGAGATGCTATGCCGAATTTTC 1440
QY 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCAAGTCTC 1500
Db 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCTCAAGTCTC 1500
QY 1501 TATTAATGAAGAACAGAGCTTGTCAAGCTCTTGTGCTGATGTCTGGAATGATG 1560
Db 1501 TATTAATGAAGAACAGAGCTTGTCAAGCTCTTGTGCTGATGTCTGGAATGATG 1560
QY 1561 TACTTCAGGAAGACATTTTCTTTTCTTGAACCTTAAGAGTCTATTAATAACAGAG 1620
Db 1561 TACTTCAGGAAGACATTTTCTTTTCTTGAACCTTAAGAGTCTATTAATAACAGAG 1620
QY 1621 ACAGATTCCACATTTTATACATGAGAGATCTTTGTGAGTGAATACAGAGATTGACTCG 1680
Db 1621 ACAGATTCCACATTTTATACATGAGAGATCTTTGTGAGTGAATACAGAGATTGACTCG 1680
QY 1681 ATCCCTTTAAAGAGCTTTATGTCCCTGACTGCTGCTAAATTTATCTAATTTCCAGATG 1740
Db 1681 ATCCCTTTAAAGAGCTTTATGTCCCTGACTGCTGCTGCTAAATTTATCTAATTTCCAGATG 1740
QY 1741 CTTTGTAGATGACTGAGATTTGTGAGGCAATATTTGGGATTTCTAGATTTGAGTGA 1800
Db 1741 CTTTGTAGATGACTGAGATTTGTGAGGCAATATTTGGGATTTCTAGATTTGAGTGA 1800
QY 1801 TGGCAGGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCTCGAAT 1860
Db 1801 TGGCAGGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCTCGAAT 1860
QY 1861 CTACAGAGAAAGAGGAGATCAGACTGAGAGAGCTGTGACATAGAGCTTGAAGCAAGA 1920
Db 1861 CTACAGAGAAAGAGGAGATCAGACTGAGAGAGCTGTGACATAGAGCTTGAAGCAAGA 1920
QY 1921 CTTTGAATTTGGAGCTGTCTATGTGTGATTAATATCACTGCTCTTCTTATATGAGT 1980
Db 1921 CTTTGAATTTGGAGCTGTCTATGTGTGATTAATATCACTGCTCTTCTTATATGAGT 1980
QY 1981 TACAAATCTATATTTTATTTGAAGTTTAAATAAGAAAAATTTACAGAAAAAATTTTAC 2040
Db 1981 TACAAATCTATATTTTATTTGAAGTTTAAATAAGAAAAATTTACAGAAAAAATTTTAC 2040
  
```

OY 2041 A 2041  
 Db 2041 A 2041  
 RESULT 2  
 LOCUS AY052768 1308 bp mRNA linear PRI 17-OCT-2001  
 DEFINITION Homo sapiens tRNA Isopentenyl transferase mRNA, partial cds.  
 ACCESSION AY052768  
 VERSION AY052768.1 GI:16209578  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1308)  
 Lemieux,J., Lakowski,B., Webb,A., Meng,Y., Ubach,A., Bussiere,F.,  
 Barnes,T. and Hekimi,S.  
 TITLE Regulation of physiological rates in *Caenorhabditis elegans* by a  
 tRNA-modifying enzyme in the mitochondria  
 JOURNAL Genetics 159 (1), 147-157 (2001)  
 MEDLINE 21444833  
 PUBMED 11560893  
 REFERENCE 2 (bases 1 to 1308)  
 Lemieux,J., Barnes,T. and Hekimi,S.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-AUG-2001) Biology, McGill University, 1205 Dr.  
 Penfield Avenue, Montreal, QC H3A 1B1, Canada  
 FEATURES  
 source location/Qualifiers  
 1..1308  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="c-2ec05"  
 <1..1308  
 /note="hgr01"  
 /codon\_start=1  
 /product="tRNA Isopentenyl transferase"  
 /protein\_id="AAL14107.1"  
 /db\_xref="GI:16209579"  
 /translation="ARAVPVSGSLGRLRLPLVVIIGATGTGKSTLALQLGRLGGE  
 IVSADSMOYVEGLDITNKVSAEORICRHHMISFVDFLVNTYVDFRNATLIED  
 IFARDKIPVGGNYIESILKLVNTKQEMGTEKVIDKYLEDEGLVLHKRL  
 SQVPEMAAKIHPHDKRVARSLQVFEFTGISHSEPLHROTEBEGSLGSLPLFSNP  
 CILWHLADQAVLDRDLRVDMLAAGLELRLDRHRTNKNVSENSQDQHGIFOS  
 IGFEFEHYLLTTECKTLETSLNQLKQGPVLPVVGLEVDYSKMEESVLEPALFI  
 VOSFIQGHKPTATPDKIMPYNEAKNRKSYHLCDLDRILIGREMAAHIKSKSHNOLK  
 KRRLDSPAVNTIESOSVSPDYNEKPEKSGSPQNDLKCYS"  
 BASE COUNT 378 a 302 c 329 g 299 t  
 ORIGIN  
 Query Match 61.6%; Score 1258; DB 9; Length 1308;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 GATCCTCTTGACCAATTACAGAGTGTGACTTCAGAAATAGAGCAACTGCTCTGATT 300  
 OY 329 GAAGTATATTTTCCCGGACAGCAAAATTCCTATTTGTGGGAGGAGCAATATATACATT 388  
 Db 301 GAAGTATATTTTCCCGGACAGCAAAATTCCTATTTGTGGGAGGAGCAATATATACATT 360  
 OY 389 GAATCTCTGCTCTGGAAGATTCTTGTCAATACCAAGCCCGAGAGATGGGCACTAGAAA 448  
 Db 361 GAATCTCTGCTCTGGAAGATTCTTGTCAATACCAAGCCCGAGAGATGGGCACTAGAAA 420  
 OY 449 GTGATTGACCGAAAGAGTGTGAAAAGAGAGATGTCTTGTATCTTACAAACGCTTA 508  
 Db 421 GTGATTGACCGAAAGAGTGTGAAAAGAGAGATGTCTTGTATCTTACAAACGCTTA 480  
 OY 509 AGCCAGTGTGACCCGAAATATGCTTCCAGCTGCATCCACATGACAAACGCAAACTGGCC 568  
 Db 481 AGCCAGTGTGACCCGAAATATGCTTCCAGCTGCATCCACATGACAAACGCAAACTGGCC 540  
 OY 569 AGGACCTTGCAAGTTTGTGAAGAAACAGAAATCTCTCATAGTAATTTCTCCATGCTCAA 628  
 Db 541 AGGACCTTGCAAGTTTGTGAAGAAACAGAAATCTCTCATAGTAATTTCTCCATGCTCAA 600  
 OY 629 CATACGGAAGAAGTGTGTGTCCTCTTGAGAGTCTCTGAAGTCTTAACCTTGATC 688  
 Db 601 CATACGGAAGAAGTGTGTGTCCTCTTGAGAGTCTCTGAAGTCTTAACCTTGATC 660  
 OY 689 CTTTGGCTTCATGCTGACCAAGGCACTTCTAGATGAGCGCTTGATTAAGAGGTGATGAC 748  
 Db 661 CTTTGGCTTCATGCTGACCAAGGCACTTCTAGATGAGCGCTTGATTAAGAGGTGATGAC 720  
 OY 749 ATGCTTGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTTCACAGACGCTAATAGAGA 808  
 Db 721 ATGCTTGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTTCACAGACGCTAATAGAGA 780  
 OY 809 AATGTTTGGAAATAGCCAGAGATATCAACATGTGTATCTTCAATTCGCTTCAAG 868  
 Db 781 AATGTTTGGAAATAGCCAGAGATATCAACATGTGTATCTTCAATTCGCTTCAAG 840  
 OY 869 GAATTTACAGATACCTGATCTACTGAGGAAATATGACACTGTGAACCTAGTACCAGCTT 928  
 Db 841 GAATTTACAGATACCTGATCTACTGAGGAAATATGACACTGTGAACCTAGTACCAGCTT 900  
 OY 929 CTAAGAAGAGACCTGTCCTGATGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 988  
 Db 901 CTAAGAAGAGACCTGTCCTGATGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTG 960  
 OY 989 TCGAAGTGGAGAGTGTGTTCTTGAACCTGCTTGAATGTCGAAGTTTCATCCAG 1048  
 Db 961 TCGAAGTGGAGAGTGTGTTCTTGAACCTGCTTGAATGTCGAAGTTTCATCCAG 1020  
 OY 1049 GGCCACAAAGCTTACAGCCACTTCAATTAAGATGCCATATGAAGCTGACACAAAGCA 1108  
 Db 1021 GGCCACAAAGCTTACAGCCACTTCAATTAAGATGCCATATGAAGCTGACACAAAGCA 1080  
 OY 1109 AATTATACCTGTGTGACCTGTGATGCAATCATTTGGGATCGGGAATGGGCACAG 1168  
 Db 1081 AATTATACCTGTGTGACCTGTGATGCAATCATTTGGGATCGGGAATGGGCACAG 1140  
 OY 1169 CACATATAATCCAAATCCCACTTGAACCACTGAAGAAAGAGAAATTGGACTCAGAT 1228  
 Db 1141 CACATATAATCCAAATCCCACTTGAACCACTGAAGAAAGAGAAATTGGACTCAGAT 1200  
 OY 1229 GCTGTCAACACCATAGAAAGTGTGATTTTCCCAAGACTATACAAAGAACCTTAAG 1286  
 Db 1201 GCTGTCAACACCATAGAAAGTGTGATTTTCCCAAGACTATACAAAGAACCTTAAG 1258  
 RESULT 3  
 AA405704 2130 bp DNA linear PAT 14-JUN-2002  
 LOCUS AA405704  
 DEFINITION Sequence 119 from Patent WO0222660.  
 ACCESSION AX405704



VERSION	AX405704.1	GI:21438847
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1	
AUTHORS	Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q. A., Ren, F., Xue, A. J., Yang, Y., Wehrman, T., and Dimaano, R. T.	
TITLE	Novel nucleic acids and polypeptides	
JOURNAL	Patent: WO 0222660-A 119 21-MAR-2002;	
	HYSEQ, INC. (US)	
FEATURES	Location/Qualifiers	
SOURCE	1..2130	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	191..1426	
CDS	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD34811.1"	
	/db_xref="GI:21438848"	
	/translation="MOYEGEDITITNKVSAQEOIRCHNMHSFVDPILVTNTVVDFFNRN	
	RATLIEDIFARDKPIYVGSTNYIISLMTKLVNTKPOEMGEKVIYDRLEVEKED	
	GLVLRKLSQVDEPMARKLPHDKRKVARSLOVEENGISHEPGLHONHEEFGPLG	
	GLPKESNPCTILHADAVIDERLDRKVDMLAGLLELDRPHRRNOKVNSNSOD	
	YQKIFOSIGFERHEYLITEGKSTLETSMOLKKGLEALQVTKRKARKONRVKRN	
	FLSHGPVIPPVYGLSEVDVSKWESEVLEPALEVSFIOGHKPTAPIKPIVYEAE	
	KRSYHLCIDLDRIIIGREMAAHIKSKSHLNQKKRRRLSDAVNTIESPDKNK	
	EPKESGSPQNDDELKCSY"	
BASE COUNT	620 a 439 c 522 g 549 t	
ORIGIN		
Query Match	43.9%; Score 896; DB 6; Length 2130;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches	896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	44 GTGGCAGTGGGCTAGGGGCGCTGCAAGGACCTTACTCTTTGTAGTAGTTCGCGGCC	103
DB	56 GTGGGCACTGGGCTTAGGGGCGCTGCAAGGACCTTACTCTTTGTAGTAGTTCGCGGCC	115
OY	104 ACGGCGACCGGCAATTCACGCTGGCGTTCGACGTAGCGCAGCGGCTGGCGGTGAGATC	163
DB	116 ACGGCGACCGGCAATTCACGCTGGCGTTCGACGTAGCGCAGCGGCTGGCGGTGAGATC	175
OY	154 GTGACGCTGACTCCATGCTAGGTCTATGAAAGGCTTAGACATCATCACCACAAGTTTCT	223
DB	176 GTGACGCTGACTCCATGCTAGGTCTATGAAAGGCTTAGACATCATCACCACAAGTTTCT	235
OY	224 GCCCAAGGACGAGAAATTCGCCGGACCAATGATGACTTTGTGATCTCTTGTGACG	283
DB	236 GCCCAAGGACGAGAAATTCGCCGGACCAATGATGACTTTGTGATCTCTTGTGACG	295
OY	284 AATTACACAGTGGTGGACTTCAGAAATAGACAACTGGCTCGATGGAAGATATATTGCC	343
DB	296 AATTACACAGTGGTGGACTTCAGAAATAGACAACTGGCTCGATGGAAGATATATTGCC	355
OY	344 CGAGCAAAATTCCTATTGTTGTGGGAGAAACAATTTATACATTGAATCTCTGCTTGG	403
DB	356 CGAGCAAAATTCCTATTGTTGTGGGAGAAACAATTTATACATTGAATCTCTGCTTGG	415
OY	404 AAATGTTCTTGCATATACCAAGCCCGACGAGATGGGCACTGAGAAAGTATTGACGAAA	463
DB	416 AAATGTTCTTGCATATACCAAGCCCGACGAGATGGGCACTGAGAAAGTATTGACGAAA	475
OY	464 GTGGAGCTTGAAGAAGAGATGCTTGTACTTCAACAAGCGCTAAGCCAGGTGAGCCA	523
DB	476 GTGGAGCTTGAAGAAGAGATGCTTGTACTTCAACAAGCGCTAAGCCAGGTGAGCCA	535
OY	524 GAAATGGCTGCAAGCTCATATCACAAGGCAAGTGGCGAGAGCTTGCAAGTT	583
DB	536 GAAATGGCTGCAAGCTCATATCACAAGGCAAGTGGCGAGAGCTTGCAAGTT	595
OY	584 TTGTGAAGAAGCAGATCTCTCATAGTGAATTTCTCATCTGCTCAACATACGGAAGAGCT	643

Db	556	TTTGTGAAAGACGAGATCTTCATAGTAATTTCTCCATCGTCACATACGGAAGAAGT	655
Oy	644	GGTGTCTCCCTTGGAGGTCCTCTGAAGTTCTCTTAACCCCTTGATCCTTTGGCTTCAGCT	703
Db	656	GGTGTCTCCCTTGGAGGTCCTCTGAAGTTCTCTTAACCCCTTGATCCTTTGGCTTCAGCT	715
Oy	704	GACCAAGCAAGTTTCTAGATGAGCGCTTGGATAAGAGGGTGGATGACATGCTTCTCTGG	763
Db	716	GACCAGGCAAGTTTCTAGATGAGCGCTTGGATAAGAGGGTGGATGACATGCTTCTCTGG	775
Oy	764	CTCTGGAGGAACCTAAGAGATTTTTCACAGAGCCCTTAATTCAGAAAGATTTTCGAAAT	823
Db	776	CTCTGGAGGAACCTAAGAGATTTTTCACAGAGCCCTTAATTCAGAAAGATTTTCGAAAT	835
Oy	824	AGCCAGGACTATCAACATGATCTTCCCAATCAATGGCTTCAAGAAATTTTCACAGATAC	883
Db	836	AGCCAGGACTATCAACATGATCTTCCCAATCAATGGCTTCAAGAAATTTTCACAGATAC	895
Oy	884	CTGATCACTGAGGGAATAATGCACACTGGAGACTAGTAACACAGCTTCTTAAGAAGG	939
Db	896	CTGATCACTGAGGGAATAATGCACACTGGAGACTAGTAACACAGCTTCTTAAGAAGG	951
RESULT 4			
AF074918			
LOCUS	AF074918	2216 bp	RNA
DEFINITION	Homo sapiens tRNA isopentenylpyrophosphate transferase precursor		
ACCESSION	AF074918		
VERSION	AF074918.1	GI:11137964	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 2216)		
AUTHORS	Golovko,A., Hjaltn,G., Sitbon,F. and Nicander,B.		
TITLE	Cloning of a human tRNA isopentenyl transferase		
JOURNAL	Gene 258 (1-2), 85-93 (2000)		
MEDLINE	20564178		
PUBMED	11111046		
REFERENCE	2 (bases 1 to 2216)		
AUTHORS	Golovko,A. and Hjaltn,G.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-JUN-1998) Dep. of Plant Biology, Swedish University		
	of Agricultural Sciences, Box 7080, SE-750 07 uppsala, Sweden		
FEATURES			
Source	1..2216		
	/organism="Homo sapiens"		
	/db_xref="dbest:AA332152"		
	/db_xref="taxon:9606"		
	/note="extension of sequence by 5' RACE"		
	1..2216		
	/note="incompletely processed mRNA"		
	join(15..829,917..1505)		
	/note="1Pase"		
	/codon_start=1		
	/product="tRNA isopentenylpyrophosphate transferase"		
	/protein_id="AA331324.1"		
	/db_xref="GI:11137965"		
	/translation="MASVAAARAVPGSGLRGLRPLVVIIGANGTGKSLALGIC		
	QRLEGISADSMQVYGLDITINRKSVAOERLICRHIMISVPLVYNYVDFRNRA		
	TALIEDFARDKIPVIGGTVYIESLMLKLVTKRQEMGTEKVIDRKLKEDGL		
	VTHRRISOVDPEMAKAPLPHDKRVARSLQVFEETGISSEFLRHRTFEGGGLGR		
	LKSNPCILMLHADAVLDELKRVYDMILACGLELRDFHRRYRNKNKNSQNDYO		
	HGIPQISGKEFHEHYLTTEGCTLETSLNLLKGIETALKOVTKRYARKONRWANRFL		
	SRPQIVPVYGLEVDVSKWESVLEPALEIVQSFLQGHKPRATPIKMPYNEAKNR		
	SYHLCDCDRILIIDREMAHIKSKSLNQLKRRRLSDAVNTIESQVSPDHKEP		
	KKESPCNDQNELKSV"		
	830..916		
	/evidence="experimental"		
	2177..2182		

BASE COUNT	651 a	461 c	534 g	570 t	ORIGIN
Query Match	40.4%	Score 825	DB 9	Length 2216	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 825	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
OY	1	CTGCCATAAGATGGCGTCCGTGGCGGCTGCAACAGACCACTTCTCTGGGGCAGTGGGCTCAG	60		
DB	5	CTGGCATTAAGATGGCGTCCGTGGCGGCTGCAACAGACCACTTCTCTGGGGCAGTGGGCTCAG	64		
OY	61	GGGGCTGCACAGGAGCCCTACCTCTTGTAGTATCTTCGGGGCCACGGGCACTGGCGAAATC	120		
DB	65	GGGGCTGCACAGGAGCCCTACCTCTTGTAGTATCTTCGGGGCCACGGGCACTGGCGAAATC	124		
OY	121	CACGCTGGCGTTTGGCAGTAGGTCACAGGGCTCGGGCGGTGAGATGTCAGCGCTGACTTCAT	180		
DB	125	CACGCTGGCGTTTGGCAGTAGGTCACAGGGCTCGGGCGGTGAGATGTCAGCGCTGACTTCAT	184		
OY	181	GCAGGCTATGAGAGGCTTAGACATCATCAACAACAAGTTTCTGCCCAAGACAGAGAAAT	240		
DB	185	GCAGGCTATGAGAGGCTTAGACATCATCAACAACAAGTTTCTGCCCAAGACAGAGAAAT	244		
OY	241	CTGCCGGACCAACATGATCAGCTTGTGTGATTCCTTGTGACCAATTCACAGTGTGGA	300		
DB	245	CTGCCGGACCAACATGATCAGCTTGTGTGATTCCTTGTGACCAATTCACAGTGTGGA	304		
OY	301	CTTCAGAATAGAGCAATCGCTCTGATTGAAGATATATTTGCCGAGACAAATTCCTAT	360		
DB	305	CTTCAGAATAGAGCAATCGCTCTGATTGAAGATATATTTGCCGAGACAAATTCCTAT	364		
OY	361	TGTTGTGGAGAGAACCAATTATTCATTTGAATCTCTGCTCTGGAAAGTTCTTGTCAATAC	420		
DB	365	TGTTGTGGAGAGAACCAATTATTCATTTGAATCTCTGCTCTGGAAAGTTCTTGTCAATAC	424		
OY	421	CAAGCCCAAGAGATGGGCACTGACAAAGTATTTGACCCAAAGTCGACCTTGAAAGGA	480		
DB	425	CAAGCCCAAGAGATGGGCACTGACAAAGTATTTGACCCAAAGTCGACCTTGAAAGGA	484		
OY	481	GGATGGTCTTGTACTTCACAAACGCCCTTAAGCCAGGTGACCCAGAAATGGTCCCAAGCT	540		
DB	485	GGATGGTCTTGTACTTCACAAACGCCCTTAAGCCAGGTGACCCAGAAATGGTCCCAAGCT	544		
OY	541	GCATCCACATGACAAACGCCAAAGTGCGCAGAGCTTGCAAGTTTTTGAAGAAACAGCAAT	600		
DB	545	GCATCCACATGACAAACGCCAAAGTGCGCAGAGCTTGCAAGTTTTTGAAGAAACAGCAAT	604		
OY	601	CTCTCATAGTGAATTTCTCCATCGTCACATACGGAAGAAGTGGTGGGCCCTTGGAGG	660		
DB	605	CTCTCATAGTGAATTTCTCCATCGTCACATACGGAAGAAGTGGTGGGCCCTTGGAGG	664		
OY	661	TCCTCTGAAGTCTCTTAACCCCTTGGATCCTTTGGCTTCATGCTGACACAGCAATCTTAGA	720		
DB	665	TCCTCTGAAGTCTCTTAACCCCTTGGATCCTTTGGCTTCATGCTGACACAGCAATCTTAGA	724		
OY	721	TGAGGCGTTGATGAAGAGGCTGATGACATGCTGCTGCTGGGCTCTTGGAGAACTAAG	780		
DB	725	TGAGGCGTTGATGAAGAGGCTGATGACATGCTGCTGCTGGGCTCTTGGAGAACTAAG	784		
OY	781	AGATTTTCACAGACGCTATATCAGAAGAAATGTTTCGGAAATAG	825		
DB	785	AGATTTTCACAGACGCTATATCAGAAGAAATGTTTCGGAAATAG	829		
RESULT 5					
LOCUS	BC010741	1844 bp	mRNA	linear	PRI 17-JUL-2001
DEFINITION	Homo sapiens, Similar to tRNA isopentenylpyrophosphate transferase, clone MGC:17002 IMAGE:3905836, mRNA, complete cds.				
ACCESSION	BC010741				
KEYWORDS	BC010741.1	GI:14789610			
SOURCE	MGC.				
	Homo sapiens.				

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 1844)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (12-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgaabs@mail.nih.gov">cgaabs@mail.nih.gov</a> Tissue Procurement: ANCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: <a href="http://www.sngc.stanford.edu">http://www.sngc.stanford.edu</a> Contact: (Dickson, Mark) <a href="mailto:mcd@pacsl.stanford.edu">mcd@pacsl.stanford.edu</a> Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Series: IRAC Plate: 14 Row: P Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 7019914.
SOURCE	Location/Qualifiers 1. 1844 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:17002 IMAGE:3905836" /tissue_type="uterus, leiomyosarcoma" /clone_id="NH_MGC_71" /lab_host="DH10B" /note="Vector: pcMV-SPORT6" 166..1140 /codon_start=1 /product="Similar to tRNA isopentenylpyrophosphate transferase" /protein_id="AAH10741.1" /db_xref="GI:14789611" /translation="MGTEKVIDRKEVELEKEDGLVLRISQVDPEMAKLPHDNRKV /ARSLQFEETGISSEFLRHROTHEEGGDLPGKFSPCLIMLHADQADERLDKRPD DMLAAGLEELRDHFRRRNOKNSSENSODYOHGIFOSIGEFHEYLITECKLETYS NOLLKGIEALKQVKTKRKARKONRWKNRFLSRPGLIIPVYGLEVDVSKESSVLE PALEIVOSFIOGHKPTAPIKMPYNEAKENRSYHLCDLDRIIIGDEMAHIKSSSH LNOKKRRRLDSDAVNITSESQSVSDHNKKEPKERKESGGONDOLKCSV"
CDS	
BASE COUNT	545 a 374 c 454 g 471 t
ORIGIN	
Query Match	37.5%; Score 765; DB 9; Length 1844;
Best Local Similarity	99.5%; Pred. No. 0;
Matches 1085:	Conservative 0; Mismatches 4; Indels 1; Gaps 1;
Dp	939 GACCTGTGCCATTGTGCCCCCCTGTCTATGCGTTAGAGGTATCTGATCTTCGAAGTGG 598       
Dp	743 GACCTGTGCCATTGTCCCCCCTGTCTATGCGTTAGAGGTATCTGATCTTCGAAGTGG 802       
Dp	999 AGGATCTGTCTTGTAACCTGCTCTTGAAATGTGTCAAAAGTTTCATCCAGGCGCACCAAG 1058       
Dp	803 AAGATCTGTCTTGTAACCTGCTCTTGAAATGTGTCAAAAGTTTCATCCAGGCGCACCAAG 862       
Dp	1059 CTACAGCCATCCCATTAAGAATGCATATCAATGAAGTGTGAGAACAGAGAAGATTATCAC 1118       
Dp	863 CTACAGCCATCCCATTAAGAATGCATATCAATGAAGTGTGAGAACAGAGAAGATTATCAC 922       
Dp	1119 TGTGTGACCTCTGTGATGGAATCATCATTTGGGGATGCGGAATGGCGACGCACATAAAT 1178       
Dp	923 TGTGTGACCTCTGTGATGGAATCATCATTTGGGGATGCGGAATGGCGACGCACATAAAT 982       
Dp	1179 CCAATGCCACTTGACCAACCATGGAAGAAAAGAGAAGATTGACATCTGCTGCACA 1238

```

Db 983 CCAATCCACCTGTAACCACTGAAGAAAGAAAGATTGGACTGAGTGTCTACACA 1042
Qy 1239 CCATGAAGAGTCAGAGTGTTCCTCCAGACTATACAAAGAACTAAAGGAGGATCCC 1298
Db 1043 CCATGAAGAGTCAGAGTGTTCCTCCAGACTATACAAAGAACTAAAGGAGGATCCC 1102
Qy 1299 CAGGAGAGATGATCAAGAGCTGAATTCAGCGCTTTAAAGACATGCTCCAGTGGCTTGG 1358
Db 1103 CAGGAGAGATGATCAAGAGCTGAATTCAGCGCTTTAAAGACATGCTCCAGTGGCTTGG 1162
Qy 1359 GAAAGTGTGTGGGATCCAGATTGAGAGGAGGGGATATGTTGTCTCCAGTGGGCAA 1418
Db 1163 GAAAGTGTGTGGGATCCAGATTGAGAGGAGGGGATATGTTGTCTCCAGTGGGCAA 1222
Qy 1419 AGGAGTGTGTGGGATCCAGATTGAGAGGAGGGGATATGTTGTCTCCAGTGGGCAA 1478
Db 1223 AGGAGTGTGTGGGATCCAGATTGAGAGGAGGGGATATGTTGTCTCCAGTGGGCAA 1282
Qy 1479 GGTTTAAAGTCTCAGCTCTCTATATAGAAACAGAGGCTGTTCAGCTCTCTGCTG 1538
Db 1283 GGTTTAAAGTCTCAGCTCTCTATATAGAAACAGAGGCTGTTCAGCTCTCTGCTG 1342
Qy 1539 GCTGATGTGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1598
Db 1343 GCTGATGTGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1402
Qy 1599 AGTGTCTATATATTAAGACGACAGATTCACATTTTATACATGAGATCTTCTTGT 1658
Db 1403 AGTGTCTATATATTAAGACGACAGATTCACATTTTATACATGAGATCTTCTTGT 1462
Qy 1659 GGTGAATACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1718
Db 1463 GGTGAATACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1521
Qy 1719 AAAATATATATATTAAGACGATGATGATGATGATGATGATGATGATGATGATGATG 1778
Db 1522 AAAATATATATTAAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1581
Qy 1779 GGGATGTTAGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1838
Db 1582 GGGATGTTAGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1641
Qy 1839 AACCAACTAGTCTGGAATTCAGAGAAAGAGGAAATGATGATGATGATGATGATGATG 1898
Db 1642 AACCAACTAGTCTGGAATTCAGAGAAAGAGGAAATGATGATGATGATGATGATGATG 1701
Qy 1899 CATAGACTTGAAGACCAAGATTTGAATTTGAGCTGCTCATGTGATGATGATGATGATG 1958
Db 1702 CATAGACTTGAAGACCAAGATTTGAATTTGAGCTGCTCATGTGATGATGATGATGATG 1761
Qy 1959 CACTGCTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2018
Db 1762 CACTGCTGTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1821
Qy 2019 AAATTACAA 2028
Db 1822 AAATTACAA 1831

```

```

RESULT 6
AK000068 1749 bp mRNA linear PRI 22-FEB-2000
DEFINITION Homo sapiens CDNA FLJ20061 fls. clone COL01383.
ACCESSION AK000068
VERSION AK000068.1 GI:7019914
KEYWORDS oligo cloning; fls (full insert sequence).
SOURCE Homo sapiens colon CDNA to mRNA, clone_11b.COL clone:COL01383.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,

```

```

FEATURES
source
    Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
    Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
    MEDO human CDNA sequencing project
    Unpublished
    2 (bases 1 to 1749)
    Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
    Shibahara,T., Tanaka,T. and Nakamura,Y.
    Direct Submission
    Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
    University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
    Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,
    Tel:81-3-5449-5286, Fax:81-3-5449-5416)
    MEDO human CDNA sequencing project supported by Ministry of
    International Trade and Industry of Japan; CDNA full insert
    sequencing; Research Association for Biotechnology; CDNA library
    construction, 5'-3'-end one pass sequencing; Department of
    Virology and Human Genome Center, Institute of Medical Science,
    University of Tokyo (partly supported by Science and Technology
    Agency).

CDS
    1..1749
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="COL01383"
    /tissue_type="colon"
    /clone_11b="COL"
    /note="cloning vector pME18SFL3"
    61..1041
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="BAA09023.1"
    /db_xref="GI:7019915"
    /translation="MTEKEVVIDRKVELEKEDGLVLRKLSQVDPMAAKLHPHDKRY
    ARSLQVEFETGISHSEFLRHQHTBEGGSLPLKFSNPLILMLHADAVIDERDKR
    VDDMLAAGLIEELRDFHRRYNGKANSSENSODVQHSIFOSIGFKEFEVLITGKCTLE
    TSNOLKKGIEALKQVTRKRYARKONRWKRNRLSPGPIVPEVGLFVSDSKMESV
    LEPALEIVSFQGHKPRATPIKMYNENKRSYHLIDICRIIIGREMAAHIKSK
    SHUNLQKRRRLSDSAVNTIESQVSPDHNKPKERKSPQNDDELKCV"

BASE COUNT
    541 a 337 c 408 g 463 t

ORIGIN
    Query Match 36.8%; Score 751; DB 9; Length 1749;
    Best Local Similarity 99.6%; Pred. No. 0;
    Matches 1021; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

QY	1421	GAGGCTATGGGGAATTCCTCGCATAGACAGAAAAGCTCCACACATTTCTTTATGATGGG	1480
Db	1126	GAGGCTATGGGGAATTCCTCGCATAGACAGAAAAGCTCCACACATTTCTTTATGATGGG	1185
OY	1481	TTTTAAAGCTCAGGCTTCCTATATATAGAAACACAGAGCTCTTGACAGCTCCTTGATGGC	1540
Db	1186	TTTTAAAGCTCAGGCTTCCTATATATAGAAACACAGAGCTCTTGACAGCTCCTTGATGGC	1245
OY	1541	TGATGTGCTGTGAAATGATGATGATTCAGAGAAAGCATTTTTTTTCTTTTGAACCTTAAG	1600
Db	1246	TGATGTGCTGTGAAATGATGATGATTCAGAGAAAGCATTTTTTTTCTTTTGAACCTTAAG	1305
OY	1601	GTTCTATTTATTTAAAGCAGCAGACATTCACATTTTATACATGAGATCTTCTTTGCG	1660
Db	1306	GTTCTATTTATTTAAAGCAGCAGATTCACATTTTATACATGAGATCTTCTTTGCG	1365
OY	1661	TGAATACACAGATTCAGATTCACATTCCTTTAAAGAAAGTTTATGTCCTGAGCTGGCTAA	1720
Db	1366	TGAATACACAGATTCAGATTCACATTCCTTTAAAGAAAGTTTATGTCCTGAGCTGGCTAA	1424
OY	1721	AATATCTAATTTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTGACACATATTGG	1780
Db	1425	AATATCTAATTTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTGACACATATTGG	1484
OY	1781	GAGTCTATGATTTGAGTAAATGGCAGGAAAGGCCATCTCCATTTAGATGATTAAGTAA	1840
Db	1485	GAGTCTATGATTTGAGTAAATGGCAGGAAAGGCCATCTCCATTTAGATGATTAAGTAA	1544
OY	1841	CCAAACTGATTCGCAATTCACAGAAAGAGGAGATTCAGACTGAGAGAGCTGTGACA	1900
Db	1545	CCAAACTGATTCGCAATTCACAGAAAGAGGAGATTCAGACTGAGAGAGCTGTGACA	1604
OY	1901	TAGAGCTTGAGACCAAGAGCTTGAATTTTGCAGATGCTCATGCTGAGTTATTATCA	1960
Db	1605	TAGAGCTTGAGACCAAGAGCTTGAATTTTGCAGATGCTCATGCTGAGTTATTATCA	1664
OY	1961	CTGCTGCTCTTCTATATGAGTACAAATCTATATTTTATGAGATTAAATPAGAAAAA	2020
Db	1665	CTGCTGCTCTTCTATATGAGTACAAATCTATATTTTATGAGATTAAATPAGAAAAA	1724
OY	2021	ATTTA 2025	
Db	1725	ATTTA 1729	
RESULT 7			
AK074222		1749 bp	linear PRI 15-FEB-2002
LOCUS	AK074222		
DEFINITION	Homo sapiens cDNA FLJ23642 fls, clone COL01371, highly similar to		
LOCUS	Homo sapiens tRNA isopentenylpyrophosphate transferase precursor		
ACCESSION	AK074222		
VERSION	AK074222.1	GI:18676765	
KEYWORDS	oligo capping; fls (full insert sequence).		
SOURCE	Homo sapiens colon cDNA to mRNA, clone, lib: COL clone: COL01371.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1		
JOURNAL	Kawabata, A., Hiki, J., Kobateke, N., Inagaki, H., Ikema, Y.,		
REFERENCE	Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,		
AUTHORS	Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T., and Sugano, S.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL	unpublished		
REFERENCE	2 (bases 1 to 1749)		
AUTHORS	Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,		
TITLE	Shibahara, T., Tanaka, T., and Nakamura, Y.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science,		
AUTHORS	University of Tokyo, Laboratory of Genome Structure, Human Genome		
TITLE	Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan		
JOURNAL	(E-mail: cdna@ims.u-tokyo.ac.jp, Tel: 81-3-5449-5286,		
REFERENCE	Fax: 81-3-5449-5416)		

COMMENT  
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan: cDNA full insert sequencing, Research Association for Biotechnology, cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES  
Source  
1. 1749  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="COL01371"  
/tissue\_type="colon"  
/clone\_1lb="COL"  
/note="cloning vector: pME18SFU3"

BASE COUNT  
541 a 337 c 408 g 463 t

ORIGIN

Query Match 36.8%; Score 751; DB 9; Length 1749;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1021; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY	1001	GAGTCGTCTTGAACCTGCCTTGAAATCGTGAAGATTTCACGAGGCCACAAAGCTT	1060
DB	706	GAGTCGTCTTGAACCTGCCTTGAAATCGTGAAGATTTCACGAGGCCACAAAGCTT	765
QY	1061	ACAGCCACTCCAAATAAGTGGCATTCATGTAAGCTGAGAAACAAGAAAGTTATCACTG	1120
DB	766	ACAGCCACTCCAAATAAGTGGCATTCATGTAAGCTGAGAAACAAGAAAGTTATCACTG	825
QY	1121	TGTGACCTGTGATGATCATCATTTGGGGATGCGGAATGGGCGACATTAATAATCC	1180
DB	826	TGTGACCTGTGATGATCATCATTTGGGGATGCGGAATGGGCGACATTAATAATCC	885
QY	1181	AAATCCCACTGACCACTGAAGAAAGAAAGAAAGATTGGACTCAGATGCTGTCAACCC	1240
DB	886	AAATCCCACTGACCACTGAAGAAAGAAAGAAAGATTGGACTCAGATGCTGTCAACCC	945
QY	1241	ATAGAAAGTCAAGATGTTTCCCCAGACTTAAACAAAGAACCTAAAGGGAAGGATCCCA	1300
DB	946	ATAGAAAGTCAAGATGTTTCCCCAGACTTAAACAAAGAACCTAAAGGGAAGGATCCCA	1005
QY	1301	GGGCGAATGATCAAGAGCTGGAATGSCACAGTTTAAAGACATGTACAGTGCCCTTGGA	1360
DB	1006	GGGCGAATGATCAAGAGCTGGAATGSCACAGTTTAAAGACATGTACAGTGCCCTTGGA	1065
QY	1361	AAGGTGCTGGGATCCAGTTTCAGAGGAGGGGTATGTTGTCCTCCACTGGGCAAG	1420
DB	1066	AAGGTGCTGGGATCCAGTTTCAGAGGAGGGGTATGTTGTCCTCCACTGGGCAAG	1125
QY	1421	GAGTGTATGCGGAATTCCTGCTAGTACGCAAGAAAGCTCCACCATTCTTTTGTATGG	1480
DB	1126	GAGTGTATGCGGAATTCCTGCTAGTACGCAAGAAAGCTCCACCATTCTTTTGTATGG	1185
QY	1481	TTTTTAAGTCTACAGTTCCTCTAATAATAGAAACAGCAGCTTGTCAGCTCCCTGTGTGC	1540
DB	1186	TTTTTAAGTCTACAGTTCCTCTAATAATAGAAACAGCAGCTTGTCAGCTCCCTGTGTGC	1245
QY	1541	TGATGCTGTGGAATGATGTAGTTCAGAGAAAGCATTTTTTTTCTTTGAACCTTAAG	1600
DB	1246	TGATGCTGTGGAATGATGTAGTTCAGAGAAAGCATTTTTTTTCTTTGAACCTTAAG	1305
QY	1601	GTTCTATTATTAAGCAGACAGATTCACATTTTTTATACATGAGAGATCTTCTTGGG	1660
DB	1306	GTTCTATTATTAAGCAGACAGATTCACATTTTTTATACATGAGAGATCTTCTTGGG	1365
QY	1661	TGAAATCCAGAGTGTGCTCATCCCTTTTAAAGAAAGTTTATGCTCCGACTCTGGTAA	1720
DB	1366	TGAAATCCAGAGTGTGCTCATCCCTTTTAAAGAAAGTTTATGCTCCGACTCTGGTAA	1424
QY	1721	AATTATCTAATTTCAGAGCTTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTGG	1780
DB	1425	AATTATCTAATTTCAGAGCTTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTGG	1484

QY	1781	GAGTTCGTGATTTGAGTAAATGGCAGAAAGGCCATCTCCATTGAGATGATTAAGTAA	1840
Db	1485	GAGTTCGTGATTTGAGTAAATGGCAGAAAGGCCATCTCCATTGAGATGATTAAGTAA	1544
QY	1841	CCAAACTAGTTCCTCGAANTCTACAGAGAGAGGGAATCAGACTGAGAGAAGCTGTGACA	1900
Db	1545	CCAAACTAGTTCCTCGAANTCTACAGAGAGAGGGAATCAGACTGAGAGAAGCTGTGACA	1604
QY	1901	TAGAGCTGAGAGACCAAGACTTTGAAATTTGGCGAGCTGCATCTGAGATTATATCA	1960
Db	1605	TAGAGCTGAGAGACCAAGACTTTGAAATTTGGCGAGCTGCATCTGAGATTATATCA	1664
QY	1961	CTGCTGCTCTTCTATATGAGTTACAAATCTATATTTTATGAGTTAAATTAAGAAAAA	2020
Db	1665	CTGCTGCTCTTCTATATGAGTTA <sup>1</sup> CAAAATCTATATTTTATGAGTTAAATTAAGAAAAA	1724
QY	2021	ATTTA 2025	
Db	1725	ATTTA 1729	

RESULT 8				
HS118J21/c				
LOCUS	166496 bp	DNA	linear	PRI_18-FEB-2000
DEFINITION	Human DNA sequence from clone RP1-118J21 on chromosome 1p34.1-35.3			

ACCESSION	AL033527
VERSION	AL033527.26
KEYWORDS	GI:6456653
SOURCE	HTG; BMB8; Cpg Island; 1-myc; morphogenetic; oncogene; osteogenic Homo sapiens.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (Pases 1 to 166496)  
Ellington, A.  
Direct Submission  
Submitted (18-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,

COMMENT

During sequence assembly data is compared from overlapping clones where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/Projects/C.elegans/wormpep> RPI-118921 is from the library RPCI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>

VECTOR: pcypac2

This sequence is the entire insert of clone RPI-118921 The true right end of clone RPI-117023 is at 3618 in this sequence. This sequence was generated from part of bacterial clone configs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/MGP/Chrl1>.

FEATURES	Location/Qualifiers
source	1. .166496
	/organism="Homo sapiens"
	/db_xref="taxon:9606"

```

/chrosome="1"
/map="p34.1-35.3"
/clone="RP1-118J21"
/clone_lib="RPC1-1"
complement(830..29381)
/gene="BMP8"
complement(join(<830..1098,3622..3732,4242..4321,
5153..5347,14838..14986,15379..15566,28682..29381))
/gene="BMP8"
/product="DJI18J21.1 (bone morphogenetic protein 8
(osteogenic protein 2))"
/notice="match: cDNAs: Em:S77477 Em:X51801 Em:M97017
Em:U39545 Em:J04566 Em:X80992 Em:M60316 Em:M60315
Em:M97016 Em:X56906 Em:AF100787 Em:U40034"
/evideance="not_experiment"
complement(join(949..1098,3622..3732,4242..4321,
complement(join(949..1098,3622..3732,4242..4321,

```

```

/note="match: proteins: SW:P22004.SW:P20722.SW:P22003
SW:P18075.TR:Q9YGH7.SW:P23359.SW:P55105.SW:P45003
SW:P34820.SW:P34821"
/codon_start=1

```

```
/evidence=not_experimental
/product="dll1872.1 (bone morphogenetic protein 8
(osteogenic protein 2))"
```

```
/protein_id="CAB75681.1"  
/db_xref="GI:7018292"  
/translation="MTALPGPIWLIALCALGGGGPGLRPDGGCPORRLGARERRDV
```

QREILAVLGLPGRRPPRAPPASRLPASADFLMLDLYHAMAGDDDEDCAPAEERLIGRA  
DLVMSFVNVERDRALGHOEPMKEFRFDLTOIPAGEATAAEFRITKVPISLHILNRT  
IUVFQETIOGROUNDECDLRLDLEFAPROGDUVLENTAARFVIRKIDKSL

NEPQANRLPGIFDDVHGSHGQVCRHRELLVSPQDILGMDWVIAPOGYSAYICEBEC

repeat\_region 1655.1748

```

/note="L2 repeat: matches 1691. .1787 of consensus"
repeat_region 1849. .2281
/note="L2 repeat: matches 2002. .2438 of consensus"

```

```
repeat_region      .2868  
2764.  
/note="FLAM_C repeat: matches 1. .133 of consensus"  
consensus
```

```
misc_feature complement(4214...4420)
/gene="BMP8"
/note="match: STS: Em:T63436"
```

```
misc_feature complement(4252, .4438)
/gene="BMP8"
/note="match: STS: Em:AA634039"
```

```
repeat_region      5972.  .6029
                    /note="MIR repeat: matches 19. .75 of consensus"
repeat_region      6203.  .6588
```

```
repeat_region      0.000  
/note="Alub repeat: matches 1. .307 of consensus"  
repeat_region      7787. .7873
```

```
misc_feature      /note="MIR repeat: matches 5/.146 of consensus"
                  complement(8369..8876)
/gene="BMP8"
```

```

/note="match: GSS: Em:B51067"
repeat_region 9109..9404
/note="Alu/b repeat: matches 1..296 of consensus"

```

repeat_region	9427	.9746	/note="AluJb repeat: matches 1. .312 of consensus"
10075	10075	10075	

```
misc_feature    10075. 10242  
                /note="match: STS: Em:241229"  
misc_feature    10626. 12048
```

```

/note="cpg island"
/evidence=not_experimental
12351. 13402
repeat region

```

```

repeat_region      /note="L2 repeat: matches 957. .2141 of consensus"
                   13474. .13591
                   /note="L1 to EBW repeat: match 179 305 of cons

```

```
repeat_region
13596. 14185
/note="LMC3 repeat: matches 7071. .7691 of consens
```

```
repeat_region 14186..14425
/note="L2 repeat: matches 2129..2417 of consensus"
```

```
repeat_region 14457..14572
/note="MIR repeat: matches 27. .147 of consensus"
repeat_region 15137..15215
/note="MIR repeat: matches 47. .137 of consensus"
repeat_region 16285..16510
/note="L2 repeat: matches 2361. .2586 of consensus"
repeat_region 17313..17597
/note="L2 repeat: matches 2445. .2745 of consensus"
repeat_region 20320..20782
/note="MIR repeat: matches 3. .541 of consensus"
repeat_region 20785..20932
/note="MIR repeat: matches 22. .175 of consensus"
repeat_region 21140..21425
/note="AlusX repeat: matches 1. .282 of consensus"
repeat_region 21927..22073
/note="MIR repeat: matches 97. .262 of consensus"
repeat_region 22185..22408
/note="L2 repeat: matches 116. .325 of consensus"
repeat_region 22409..22704
/note="AlusX repeat: matches 5. .306 of consensus"
repeat_region 22705..22759
/note="L2 repeat: matches 60. .116 of consensus"
repeat_region 22829..22898
/note="MIR repeat: matches 5. .71 of consensus"
repeat_region 22952..22989
/note="MIR repeat: matches 227. .261 of consensus"
repeat_region 22990..23298
/note="AlusX repeat: matches 1. .308 of consensus"
repeat_region 23299..23381
/note="MIR repeat: matches 132. .227 of consensus"
repeat_region 23401..23576
/note="AlusX repeat: matches 51. .223 of consensus"
repeat_region 23600..23742
/note="MIR repeat: matches 8. .153 of consensus"
repeat_region 24157..24451
/note="AlusX repeat: matches 2. .294 of consensus"
repeat_region 24618..24684
/note="MIR repeat: matches 80. .147 of consensus"
repeat_region 24753..24891
/note="MIR repeat: matches 50. .199 of consensus"
repeat_region 25018..25159
/note="L2 repeat: matches 2606. .2750 of consensus"
repeat_region 26215..26275
/note="AlusX repeat: matches 143. .195 of consensus"
repeat_region 26504..26643
/note="MIR repeat: matches 80. .119 of consensus"
repeat_region 26930..27021
/note="MIR repeat: matches 81. .179 of consensus"
repeat_region 27580..27733
/note="MIR repeat: matches 37. .197 of consensus"
repeat_region 28059..28151
/note="L2 repeat: matches 2598. .2703 of consensus"
misc_feature 28542..30030
/note="Cpg island"
evidence="not_experimental"
repeat_region 29451..29584
/note="67 copies 2 mer 9g 59% conserved"
repeat_region 30126..30262
/note="L2 repeat: matches 2350. .2475 of consensus"
repeat_region 30283..30460
/note="MIR repeat: matches 60. .241 of consensus"
repeat_region 30928..31015
/note="MIR repeat: matches 81. .165 of consensus"
repeat_region 31389..31662
/note="AlusX repeat: matches 19. .300 of consensus"
repeat_region 31873..32143
/note="L2 repeat: matches 1468. .1747 of consensus"
repeat_region 32545..32588
/note="22 copies 2 mer aa 75% conserved"
repeat_region 32716..32751
/note="L2 repeat: matches 2244. .2278 of consensus"
repeat_region 32752..33074
/note="AluYb8 repeat: matches 1. .317 of consensus"
```

```
repeat_region 33075..33487
/note="L2 repeat: matches 2278. .2701 of consensus"
repeat_region 33864..34154
/note="AlusX repeat: matches 3. .294 of consensus"

Query Match 29.1% Score 593; DB 9; Length 166496;
Best Local Similarity 99.5%; Pred. No. 5.7e-313;
Matches 863; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1164 CAGCGCATATAATCCCAATCCCACTGAACCACTGAAGAAAGAAAGATTGACT 1223
DB 82446 CAGCGCATATAATCCCAATCCCACTGAACCACTGAAGAAAGAAAGATTGACT 82387
QY 1224 CAGATGCTGTCAACACCATATAGAAAGTCAAGTGTTCCTCCAGACTATTAACAAGACCTA 1283
DB 82386 CAGATGCTGTCAACACCATATAGAAAGTCAAGTGTTCCTCCAGACTATTAACAAGACCTA 82327
QY 1284 AAGGAGAGGATCCCCAGGAGAGATGATCAAGAGCTAAATGACGCTTTACAGACAT 1343
DB 82326 AAGGAGAGGATCCCCAGGAGAGATGATCAAGAGCTAAATGACGCTTTACAGACAT 82267
QY 1344 GTCCAGTGGCCTTTGAAAGGATGGGATCCAGTTCAGGAGGAGGATGTTGTC 1403
DB 82266 GTCCAGTGGCCTTTGAAAGGATGGGATCCAGTTCAGGAGGAGGATGTTGTC 82207
QY 1404 TCCAGTGTGGCAAGAGAGTGTATGCGAATTTCTGATAGCAGAAAGCTCCACC 1463
DB 82206 TCCAGTGTGGCAAGAGAGTGTATGCGAATTTCTGATAGCAGAAAGCTCCACC 82147
QY 1464 ATTTCTTTTGTATGTGTTTAAAGTCTCAGCTTCTATATATAGAAACAGCAGCTTTG 1523
DB 82146 ATTTCTTTTGTATGTGTTTAAAGTCTCAGCTTCTATATATAGAAACAGCAGCTTTG 82087
QY 1524 TCAGCTCTTGTGTGCTGATGTCTGGAATGATGATGATCAGAAAGCATTTT 1583
DB 82086 TCAGCTCTTGTGTGCTGATGTCTGGAATGATGATGATCAGAAAGCATTTT 82027
QY 1584 TTCTTTGAACCTTAAGGTTCTATATATTAAGAGCAGACAGATTCACATTTTATACAT 1643
DB 82026 TTCTTTGAACCTTAAGGTTCTATATATTAAGAGCAGACAGATTCACATTTTATACAT 81967
QY 1644 GAGGATCTTCTTGTGTGATATACAGAGATTCATCCCTTTAAAGAGTTTATG 1703
DB 81966 GAGGATCTTCTTGTGTGATATACAGAGATTCATCCCTTTAAAGAGTTTATG 81908
QY 1704 TCCCTGACATGCGCAAAATATATCAATTCAGATGCTTGTGATGATGATGAT 1763
DB 81907 TCCCTGACATGCGCAAAATATATCAATTCAGATGCTTGTGATGATGATGATGAT 81848
QY 1764 TGTGAGCCACATATTGGAGTCTAGATTGATGATGATGATGATGATGATGAT 1823
DB 81847 TGTGAGCCACATATTGGAGTCTAGATTGATGATGATGATGATGATGATGAT 81788
QY 1824 TGAGATGATTAAAGTGAACCAATAGTTCTCGGAATTTACAGAGAGAGGAAATCGA 1883
DB 81787 TGAGATGATTAAAGTGAACCAATAGTTCTCGGAATTTACAGAGAGAGGAAATCGA 81728
QY 1884 CTGAGAGACCTGTGACATAGACTGGAAGCAAGACATTTAAATTTGCGAGCTCA 1943
DB 81727 CTGAGAGACCTGTGACATAGACTGGAAGCAAGACATTTAAATTTGCGAGCTCA 81668
QY 1944 TGTGTGAGTTATTATACAGCTGTCTTTCTATTTGATTAACAATCTATATTTTATGAA 2003
DB 81667 TGTGTGAGTTATTATACAGCTGTCTTTCTATTTGATTAACAATCTATATTTTATGAA 81608

QY 2004 GTTAAATTAAGAAAAATTTTACAGA 2030
DB 81607 GTTAAATTAAGAAAAATTTTACAGA 81581

RESULT 9
AC025002/c AC025002 200484 bp DNA linear HTG 14-JUN-2000
LOCUS
```

```

DEFINITION Homo sapiens chromosome 1 clone RP11-204L3 map 1, *** SEQUENCING IN
PROGRESS ***. 81 unordered pieces.
ACCESSION AC025002
VERSION AC025002.3 GI:8516086
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 200484)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 1, clone RP11-204L3
REFERENCE 2 (bases 1 to 200484)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,X., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Coyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teague,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 14, 2000 this sequence version replaced gi:7770499.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7675
Center clone name: 204_L_3
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 81 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1048: contig of 1048 bp in length
* 1049 1148: gap of 100 bp
* 1149 2470: contig of 1322 bp in length
* 2471 2570: gap of 100 bp
* 2571 3643: contig of 1073 bp in length
* 3644 3743: gap of 100 bp
* 3744 4980: contig of 1237 bp in length
* 4981 5080: gap of 100 bp
* 5081 6352: contig of 1272 bp in length
* 6353 6452: gap of 100 bp
* 6453 7549: contig of 1097 bp in length
* 7550 7649: gap of 100 bp
* 7650 8977: contig of 1328 bp in length
* 8978 9077: gap of 100 bp
*
* 9078 10103: contig of 1026 bp in length
* 10104 10203: gap of 100 bp
* 10204 11362: contig of 1159 bp in length
* 11363 11462: gap of 100 bp
* 11463 12821: contig of 1359 bp in length
* 12822 12921: gap of 100 bp
* 12922 14007: contig of 1086 bp in length
* 14008 14107: gap of 100 bp
* 14108 15283: contig of 1176 bp in length
* 15284 15383: gap of 100 bp
* 15384 16399: contig of 1016 bp in length
* 16400 16499: gap of 100 bp
* 16500 17804: contig of 1305 bp in length
* 17805 17904: gap of 100 bp
* 17905 19176: contig of 1272 bp in length
* 19177 19276: gap of 100 bp
* 19277 20384: contig of 1108 bp in length
* 20385 20484: gap of 100 bp
* 20485 21689: contig of 1205 bp in length
* 21690 21789: gap of 100 bp
* 21790 23191: contig of 1402 bp in length
* 23192 23291: gap of 100 bp
* 23292 24513: contig of 1222 bp in length
* 24514 24613: gap of 100 bp
* 24614 26135: contig of 1522 bp in length
* 26136 26235: gap of 100 bp
* 26236 27373: contig of 1138 bp in length
* 27374 27473: gap of 100 bp
* 27474 28723: contig of 1250 bp in length
* 28724 28823: gap of 100 bp
* 28824 30217: contig of 1394 bp in length
* 30218 30317: gap of 100 bp
* 30318 32056: contig of 1739 bp in length
* 32057 32156: gap of 100 bp
* 32157 33818: contig of 1662 bp in length
* 33819 33918: gap of 100 bp
* 33919 35156: contig of 1238 bp in length
* 35157 35256: gap of 100 bp
* 35257 36712: contig of 1456 bp in length
* 36713 36812: gap of 100 bp
* 36813 38379: contig of 1567 bp in length
* 38380 38479: gap of 100 bp
* 38480 40190: contig of 1711 bp in length
* 40191 40290: gap of 100 bp
* 40291 41760: contig of 1490 bp in length
* 41761 41880: gap of 100 bp
* 41881 44053: contig of 2173 bp in length
* 44054 44153: gap of 100 bp
* 44154 45474: contig of 1321 bp in length
* 45475 45574: gap of 100 bp
* 45575 46821: contig of 1247 bp in length
* 46822 46921: gap of 100 bp
* 46922 47302: contig of 381 bp in length
* 47303 47402: gap of 100 bp
* 47403 49323: contig of 1921 bp in length
* 49324 49423: gap of 100 bp
* 49424 51140: contig of 1717 bp in length
* 51141 51240: gap of 100 bp
* 51241 53099: contig of 1859 bp in length
* 53100 53199: gap of 100 bp
* 53200 54631: contig of 1432 bp in length
* 54632 54731: gap of 100 bp
* 54732 55751: contig of 1020 bp in length
* 55752 55851: gap of 100 bp
* 55852 57470: contig of 1619 bp in length
* 57471 57570: gap of 100 bp
* 57571 60182: contig of 2612 bp in length
* 60183 60282: gap of 100 bp
* 60283 63065: contig of 2783 bp in length
* 63066 63165: gap of 100 bp
* 63166 64443: contig of 1278 bp in length
* 64444 64543: gap of 100 bp
* 64544 66043: contig of 1500 bp in length

```

```

* 66044 66143: gap of 100 bp
* 66144 68623: contig of 2480 bp in length
* 68624 68723: gap of 100 bp
* 68724 70122: contig of 1399 bp in length
* 70123 70222: gap of 100 bp
* 70223 72407: contig of 2185 bp in length
* 72408 72507: gap of 100 bp
* 72508 74635: contig of 2128 bp in length
* 74636 74735: gap of 100 bp
* 74736 77554: contig of 2819 bp in length
* 77555 77654: gap of 100 bp
* 77655 80277: contig of 2623 bp in length
* 80278 80377: gap of 100 bp
* 80378 81798: contig of 1421 bp in length
* 81799 81898: gap of 100 bp
* 81899 84126: contig of 2228 bp in length
* 84127 84226: gap of 100 bp
* 84227 86794: contig of 2568 bp in length
* 86795 86894: gap of 100 bp
* 86895 90311: contig of 3417 bp in length
* 90312 90411: gap of 100 bp
* 90412 92255: contig of 1844 bp in length
* 92256 92355: gap of 100 bp
* 92356 94570: contig of 2215 bp in length
* 94571 94670: gap of 100 bp
* 94671 97163: contig of 2493 bp in length
* 97164 97263: gap of 100 bp
* 97264 99724: contig of 2461 bp in length
* 99725 99824: gap of 100 bp
* 99825 103151: contig of 3327 bp in length
* 103152 103251: gap of 100 bp
* 103252 106370: contig of 3119 bp in length
* 106371 106470: gap of 100 bp
* 106471 108888: contig of 2418 bp in length
* 108889 108988: gap of 100 bp
* 108989 111514: contig of 2526 bp in length
* 111515 111614: gap of 100 bp
* 111615 114360: contig of 2746 bp in length
* 114361 114460: gap of 100 bp
* 114461 118269: contig of 3809 bp in length
* 118270 118369: gap of 100 bp
* 118370 122359: contig of 3990 bp in length
* 122360 122459: gap of 100 bp
* 122460 125669: contig of 3210 bp in length
* 125670 125769: gap of 100 bp
* 125770 129306: contig of 3537 bp in length
* 129307 129406: gap of 100 bp
* 129407 132984: contig of 3578 bp in length
* 132985 133084: gap of 100 bp
* 133085 137360: contig of 4276 bp in length
* 137361 137460: gap of 100 bp
* 137461 140476: contig of 3016 bp in length
* 140477 140576: gap of 100 bp

```

Query Match 29.1% Score 593: DB 2: Length 200484;  
 Best Local Similarity 99.5%: Pred. No. 5.7e-313;  
 Matches 863: Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

Qy 1164 CAGCCACATTAATCAATCCACTTGAACCACTGAAGAAAAGAGAATGGACT 1223
Db 94028 CAGCCACATTAATCAATCCACTTGAACCACTGAAGAAAAGAGAATGGACT 93969
Qy 1224 CAGATGCTGTCAACACCATAGAAAGTCAAGTGTTCCTCCACACTATAACAAAGAACCTA 1283
Db 93968 CAGATGCTGTCAACACCATAGAAAGTCAAGTGTTCCTCCACACTATAACAAAGAACCTA 93909
Qy 1284 AAGGGAAGGATCCCGAGGCAAGATGATCAAGAGCTGAATGACACGTTTAAGACACAT 1343
Db 93908 AAGGGAAGGATCCCGAGGCAAGATGATCAAGAGCTGAATGACACGTTTAAGACACAT 93849
Qy 1344 GTCCAGTGGCCTTTGAAAGTGTGGGGATCCAGTTCAAGAGGAGGAGGTATGTTTGC 1403
Db 93848 GTCCAGTGGCCTTTGAAAGTGTGGGGATCCAGTTCAAGAGGAGGAGGTATGTTTGC 93789

```

```

Qy 1404 TCCGAGTCTGGGCAAGAGATGCTATGCGGAATTCCTGATAGCAGAAAAGCTCCACC 1463
Db 93788 TCCGAGTCTGGGCAAGAGATGCTATGCGGAATTCCTGATAGCAGAAAAGCTCCACC 93729
Qy 1464 ATTTCTTTTGTGATGTTTAAAGTCTACGTTCTCTATATATACAAACAGCAGTCTTG 1523
Db 93728 ATTTCTTTTGTGATGTTTAAAGTCTACGTTCTCTATATATATACAAACAGCAGTCTTG 93669
Qy 1524 TCAGTCTTGTGGGCGATGATGCTGGAATATGATAGTATGATCAGAAAAGCTTTT 1583
Db 93668 TCAGTCTTGTGGGCGATGATGCTGGAATATGATAGTATGATCAGAAAAGCTTTT 93609
Qy 1584 TTCTTTGAACCTTAAGGTTCTATATTAATAAGCAGACACAGATTCACATTTTATACAT 1643
Db 93608 TTCTTTGAACCTTAAGGTTCTATATTAATAAGCAGACACAGATTCACATTTTATACAT 93549
Qy 1644 GAGGATCTTCTTTGTGGGTAATACAGAGATTTACTGTCATCCCTTTAAAAAGAGTTTATG 1703
Db 93548 GAGGATCTTCTTTGTGGGTAATACAGAGATTTACTGTCATCCCTTT AAAAGAGTTTATG 93490
Qy 1704 TCCGAGCTCTGGCTAAATTTCTAATTTCCAGATGCTTTGTGATGATCAGATAT 1763
Db 93489 TCCGAGCTCTGGCTAAATTTCTAATTTCCAGATGCTTTGTGATGATCAGATAT 93430
Qy 1764 TGTGAGCCACATATTGGAGTCTAGATTGATGATGAGGAGGAGGAGGATCCAT 1823
Db 93429 TGTGAGCCACATATTGGAGTCTAGATTGATGATGAGGAGGAGGAGGATCCAT 93370
Qy 1824 TGAGATGATTAAAGTAACCAACTAGTCTCGGAATTTACAGAGAGGAGGATACAG 1883
Db 93369 TGAGATGATTAAAGTAACCAACTAGTCTCGGAATTTACAGAGAGGAGGATACAG 93310
Qy 1884 CTGAGAGAGCTGTGACATGAGACTGGAACCAAGACATTTAAATTTGCGAGTCTCA 1943
Db 93309 CTGAGAGAGCTGTGACATGAGACTGGAACCAAGACATTTAAATTTGCGAGTCTCA 93250
Qy 1944 TGTGAGATTATATACAGCTGCTGTTTATGATGATTAATTTTATTTTAA 2003
Db 93249 TGTGAGATTATATACAGCTGCTGTTTATGATGATTAATTTTATTTTAA 93190
Qy 2004 GTTTAAATTAAGAAAAATTTACAGA 2030
Db 93189 GTTTAAATTAAGAAAAATTTACAGA 93163

RESULT 10
AL354888 199863 bp DNA linear HTG 13-JUN-2001
LOCUS AL354888/C
DEFINITION Homo sapiens chromosome 1 clone RP11-473P22, V* SEQUENCING IN
PROGRESS ***, 16 unordered pieces.
ACCESSION AL354888
VERSION AL354888.6 GI:9863692
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
Plumb/B.
REFERENCE 1 (bases 1 to 199863)
AUTHORS Direct Submission
TITLE Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
JOURNAL CB10 15A, UK. E-mail enquiries: humquery@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:921379.
COMMENT
----- Sanger Centre
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: ba473p22
----- Summary Statistics

```



```
Assembly program: XCAP4: version 4.5
Sequencing vector: plasmid: 108752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 189476 bases at least Q40
Consensus quality: 193311 bases at least Q30
Consensus quality: 195510 bases at least Q20
Insert size: 198363; sum-of-contigs
Insert size: 201159; 6.2% error; agarose-ff
Quality coverage: 3.87x in Q20 bases; sum-of-contigs Quality
Coverage: 3.98x in Q20 bases; agarose-ff
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
2278: contig of 2278 bp in length
2279: gap of 100 bp
2379: 22743: contig of 20365 bp in length
22744: 22843: gap of 100 bp
22844: 45183: contig of 22340 bp in length
45184: 45283: gap of 100 bp
45284: 49581: contig of 4298 bp in length
49582: 49681: gap of 100 bp
49682: 58258: contig of 8577 bp in length
58259: 58358: gap of 100 bp
58359: 64082: contig of 5724 bp in length
64083: 64182: gap of 100 bp
64183: 72764: contig of 8582 bp in length
72765: 72864: gap of 100 bp
72865: 86552: contig of 13688 bp in length
86553: 86652: gap of 100 bp
86653: 93481: contig of 6829 bp in length
93482: 93581: gap of 100 bp
93582: 101614: contig of 8033 bp in length
101615: 101714: gap of 100 bp
101715: 107836: contig of 6122 bp in length
107837: 107936: gap of 100 bp
107937: 144546: contig of 36610 bp in length
144547: 144646: gap of 100 bp
144647: 155892: contig of 11246 bp in length
155893: 155992: gap of 100 bp
155993: 167875: contig of 11883 bp in length
167876: 167975: gap of 100 bp
167976: 183411: contig of 15436 bp in length
183412: 183511: gap of 100 bp
183512: 199863: contig of 16352 bp in length.
Location/Qualifiers
1..199863
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-473P22"
/clone_1lb="RP11-11.2"
1..2278
/feature="assembly-fragment:00765
fragment_chain:1"
clone_end:17
vector_side:left"
2379..22743
/feature="assembly-fragment:01823
fragment_chain:1"
22844..45183
/feature="assembly-fragment:01809
fragment_chain:2"
45284..49581
/feature="assembly-fragment:01177
fragment_chain:2"
49682..58258
/feature="assembly-fragment:00999
```

```
fragment_chain:2"
58359..64082
/feature="assembly-fragment:01187
fragment_chain:2"
64183..72764
/feature="assembly-fragment:00292
fragment_chain:2"
72865..86552
/feature="assembly-fragment:00339
fragment_chain:3"
86653..93481
/feature="assembly-fragment:01156
fragment_chain:3"
93582..101614
/feature="assembly-fragment:01960
fragment_chain:3"
101715..107836
/feature="assembly-fragment:00867
fragment_chain:4"
107937..144546
/feature="assembly-fragment:01896
fragment_chain:4"
144647..155892
/feature="assembly-fragment:00307"
155993..167875
/feature="assembly-fragment:00528"
167976..183411
/feature="assembly-fragment:01170"
183512..199863
/feature="assembly-fragment:01894"
/feature="assembly-fragment:00999"
BASE COUNT 53282 a 45915 c 45981 g 53154 t 1531 others
ORIGIN
Query Match 23.2%; Score 473; DB 2; Length 199863;
Best Local Similarity 99.4%; Pred. No. 3.4e-247;
Matches 863; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 1164 CACCGACATTAATCAATCCACTTGAACCACTGAAGAAGAAGAGATTGGACT 1223
DB 26141 CACCGACATTAATCAATCCACTTGAACCACTGAAGAAGAAGAGATTGGACT 26082
QY 1224 CAGATCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTATAACAAGAACTTA 1283
DB 26081 CAGATCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTATAACAAGAACTTA 26022
QY 1284 AAGGGAAGGATATCCCAAGGACCAATATGATCAAGAGTGAATTCACACCTTTAAGACAT 1343
DB 26021 AAGGGAAGGATATCCCAAGGACCAATATGATCAAGAGTGAATTCACACCTTTAAGACAT 25962
QY 1344 GTCCAGTGGCTTTGGAAGAGTGTGGGATCCAGTTCAAGAGGAGGATATGTTGTC 1403
DB 25961 GTCCAGTGGCTTTGGAAGAGTGTGGGATCCAGTTCAAGAGGAGGATATGTTGTC 25902
QY 1404 TCCCACTTGGGCAAGAGATGCTATGCGGAATTCCTGCTATAGCAAGAAAGCTCCACC 1463
DB 25901 TCCCACTTGGGCAAGAGATGCTATGCGGAATTCCTGCTATAGCAAGAAAGCTCCACC 25842
QY 1464 ATTTCTTTTGAATGTGTTTAAAGTCTACGTTCTCTATATAAAGAACAGCAGGCTTG 1523
DB 25841 ATTTCTTTTGAATGTGTTTAAAGTCTACGTTCTCTATATAAAGAACAGCAGGCTTG 25782
QY 1524 TCAGCTCCTTGTGGTGTGCTGTGCTGGAAGATGATGTTAGAGAAAGCATTTTGT 1583
DB 25781 TCAGCTCCTTGTGGTGTGCTGTGCTGGAAGATGATGTTAGAGAAAGCATTTTGT 25722
QY 1584 TTCTTTGAACCTTAAGGTTCTATTATTAAGACAGACAGATTCACA-TTTTATPACA 1642
DB 25721 TTCTTTGAACCTTAAGGTTCTATTATTAAGACAGACAGATTCACA-TTTTATPACA 25662
QY 1643 TGAGGATCTTCTTGTGTGTAATACAGAGATGATGATCCCTTTAAAGAAAGTTTAT 1702
DB 25661 TGAGGATCTTCTTGTGTGTAATACAGAGATGATGATCCCTTT-AAAGAGTTTAT 25603
```

QY 1703 GTCCCTGACTCTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAGATGCTGAAGTAT 1762  
 |||||||  
 Db 25602 GTCCCTGACTCTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAGATGCTGAAGTAT 25543  
 |||||||  
 QY 1763 TTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTANTGGCAGAAAGGCCATCTCCA 1822  
 |||||||  
 Db 25542 TTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTANTGGCAGAAAGGCCATCTCCA 25483  
 |||||||  
 QY 1823 TTGAGATGATTAGTGAACCAACTGTTCTCGGAATTCTACAGAGAGAGGAATCAG 1882  
 |||||||  
 Db 25482 TTGAGATGATTAGTGAACCAACTGTTCTCGGAATTCTACAGAGAGAGGAATCAG 25423  
 |||||||  
 QY 1883 ACTGAGAGACCTGTGACATATGAGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTC 1942  
 |||||||  
 Db 25422 ACTGAGAGACCTGTGACATATGAGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTC 25363  
 |||||||  
 QY 1943 ATGTGTGAGTATTATTCACGTGCTCTTCTTCTATTGAGTACCAATCTATATTTTATTGA 2002  
 |||||||  
 Db 25362 ATGTGTGAGTATTATTCACGTGCTCTTCTTCTATTGAGTACCAATCTATATTTTATTGA 25303  
 |||||||  
 QY 2003 AGTTTAAATTAAGAAAAATTTACAGA 2030  
 |||||||  
 Db 25302 AGTTTAAATTAAGAAAAATTTACAGA 25275  
 |||||||

## RESULT 11

AX336538/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 341: Conservative

11.8%; Score 241; DB 6; Length 355;

99.4%; Pred. No. 6.8e-120;

0; Mismatches 2; Indels 0; Gaps 0;

QY 1690 AAAGAACTTTATGTCCTGCTGCTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAG 1749  
 |||||||  
 Db 343 AAAGAACTTTATGTCCTGCTGCTGCTAAATTAATCTAATTTCCAGATGCTTTTGTAG 284  
 |||||||  
 QY 1750 ATGACGATGATTTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTGAATGCGAGAA 1809  
 |||||||  
 Db 283 ATGACGATGATTTGTGAGCCACATATTGGAGCTTCTAGATTGTAGTGAATGCGAGAA 224  
 |||||||  
 QY 1810 AGGCCCATCTTCATTTGAGATGATTAAGTGAACCAACTAGTTCGCGAATTTCTACAGAA 1869  
 |||||||  
 Db 223 AGGCCCATCTTCATTTGAGATGATTAAGTGAACCAACTAGTTCGCGAATTTCTACAGAA 164  
 |||||||  
 QY 1870 AGGAGGAATCAGACTGAGAGAGCTGTGACATAGACTTGAAGCAACCAAGACTTTGAAT 1929  
 |||||||  
 Db 163 AGGAGGAATCAGACTGAGAGAGCTGTGACATAGACTTGAAGCAACCAAGACTTTGAAT 104  
 |||||||  
 QY 1930 TTGGGAGCTGCTCATGTGTGAGTATTATATCACTGCTCTTTCTATTGAGTTACAAATCT 1989  
 |||||||  
 Db 103 TTGGGAGCTGCTCATGTGTGAGTATTATATCACTGCTCTTTCTATTGAGTTACAAATCT 44  
 |||||||

QY 1990 ATATTTTATTGAGTTTAAATTAAGAAAAATTTACAGAAA 2032  
 |||||||  
 Db 43 ATATTTTATTGAGTTTAAATTAAGAAAAATTTACAGAAA 1

## RESULT 12

G24438/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

G24438 252 bp DNA linear STS 31-MAY-1996  
 human STS WI-12773, sequence tagged site.

G24438.1 GI:1344764

STS: STS sequence; primer: sequence tagged site.

Homo sapiens STS derived from sequences in dbEST and the Unigene collection.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 252)

Whitehead Institute/MIT Center for Genome Research; Physically

Unpublished (1995)

Contact: Thomas Hudson

Whitehead Institute/MIT Center for Genome Research

Whitehead Institute for Biomedical Research

9 Cambridge Center, Cambridge MA 02142 USA

Tel: 617 252 1900

Fax: 617 252 1902

Email: thudson@genome.wi.mit.edu

Primer A: TAATCTCATAGAAAGACGACGCTGA

Primer B: AGTTCTAGATTGAGTGAATGCA

STS size: 201

PCR Profile:

Presoak:

Denaturation:

Annealing: 56 degrees C

Polymerization:

PCR Cycles: 35

Thermal Cycler:

Protocol:

Template: 10 ng

Primer: each 5 pm

dNTPs: each 4 mM

Tag Polymerase: 0.025 units/ul

Total Vol: 20 ul

Buffer:

MgCl2: 1.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 9.3

Derived from dbEST (genbank accession Z40724).

Location/Qualifiers

1. 252

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/map="122.8 CR from top of Chr1 linkage group"

STS

primer\_bind

primer\_bind

complement(225..248)

BASE COUNT

74 a 58 c 34 g 85 t 1 others

## Query Match

Best Local Similarity

Matches 250: Conservative

7.3%; Score 150; DB 11; Length 252;

99.2%; Pred. No. 5e-70;

0; Mismatches 2; Indels 0; Gaps 0;

QY 1778 TGGGAGTTGAGATTGAGTGAATGCGAAGAGGCCATCTCCATTGAGATGATTAAGT 1837  
 |||||||

*	49682	58258:	contig of 8577 bp	in length
*	58259	58358:	gap of 100 bp	
*	58359	64082:	contig of 5724 bp	in length
*	64083	64182:	gap of 100 bp	
*	64183	72764:	contig of 8562 bp	in length
*	72765	72864:	gap of 100 bp	
*	72865	86552:	contig of 13688 bp	in length
*	86553	86652:	gap of 100 bp	
*	86653	93481:	contig of 6829 bp	in length
*	93482	93581:	gap of 100 bp	
*	93582	101614:	contig of 8033 bp	in length
*	101615	101714:	gap of 100 bp	
*	101715	107836:	contig of 6122 bp	in length
*	107837	107936:	gap of 100 bp	
*	107937	144546:	contig of 36610 bp	in length
*	144547	144646:	gap of 100 bp	
*	144647	155892:	contig of 11246 bp	in length
*	155893	155992:	gap of 100 bp	
*	155993	167875:	contig of 11883 bp	in length
*	167876	167975:	gap of 100 bp	
*	167976	183411:	contig of 15436 bp	in length
*	183412	183511:	gap of 100 bp	
*	183512	199863:	contig of 16352 bp	in length
FEATURES				
source		Location/Qualifiers		
		1..199863		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/chromosome="1"		
		/clone="RP11-473P22"		
		/clone_1fb="RPC1-11.2"		
		1..2278		
misc_feature		/note="assembly_fragment:00765		
		fragment_chain:1		
misc_feature		clone_end:"7		
		vector_side:left"		
		2379..22743		
misc_feature		/note="assembly_fragment:01823		
		fragment_chain:1"		
		22844..45183		
misc_feature		/note="assembly_fragment:01809		
		fragment_chain:2"		
		45284..49581		
misc_feature		/note="assembly_fragment:01177		
		fragment_chain:2"		
		49682..58258		
misc_feature		/note="assembly_fragment:00999		
		fragment_chain:2"		
		58359..64082		
misc_feature		/note="assembly_fragment:01187		
		fragment_chain:2"		
		64183..72764		
misc_feature		/note="assembly_fragment:00292		
		fragment_chain:2"		
		72865..86552		
misc_feature		/note="assembly_fragment:00339		
		fragment_chain:3"		
		86553..93481		
misc_feature		/note="assembly_fragment:01156		
		fragment_chain:3"		
		93582..101614		
misc_feature		/note="assembly_fragment:01960		
		fragment_chain:3"		
		101715..107836		
misc_feature		/note="assembly_fragment:00867		
		fragment_chain:4"		
		107937..144546		
misc_feature		/note="assembly_fragment:01896		
		fragment_chain:4"		
		144647..155892		
misc_feature		/note="assembly_fragment:00307"		
		155993..167875		
misc_feature		/note="assembly_fragment:00528"		
		167976..183411		

```

/note="assembly_fragment:01170"
misc.feature      /note="assembly_fragment:01894"
BASE COUNT       53282 a 45915 c 45981 g 53154 t 1531 others
ORIGIN
Query Match      7 0%: Score 143; DB 2; Length 199863;
Best Local Similarity 100.0%; Pred. No. 2,1e-66;
Matches 143: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 44 GTGGGAGTGGGCTCAGGGGCTGCAAGGACCTACTCTTGTAGTATTCCTGGGGCC 103
DB 8752 GTGGGAGTGGGCTCAGGGGCTGCAAGGACCTACTCTTGTAGTATTCCTGGGGCC 8811

Y 104 ACGGGACCGGCAATCACTGCTGCTGACCTAGGCCAGCGGCTGGCGGTGAGATC 163
DB 8812 ACGGGACCGGCAATCACTGCTGCTGACCTAGGCCAGCGGCTGGCGGTGAGATC 8871

Y 164 GTGAGCGCTGACTCCATGACAGCT 186
DB 8872 GTGAGCGCTGACTCCATGACAGCT 8894

RESULT 14
AC025068      79785 bp      DNA      linear      HTG 13-JUL-2000
LOCUS         AC025068/c
DEFINITION    Homo sapiens chromosome 1 clone RP11-21702 map 1, LOW-PASS SEQUENCE
ACCESSION     AC025068
VERSION       AC025068.2 GI:7329438
KEYWORDS      HTG; HTGS_PHASE0.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
              1 (bases 1 to 79785)
              Homo sapiens chromosome 1, clone RP11-21702
              Unpublished
              2 (bases 1 to 79785)
REFERENCE     Biren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
              Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
              Bonuslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
              Campoliano, A., Castelle, A., Choquel, Y., Colangelo, M., Collins, S.,
              Collamore, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S.,
              Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
              Galagan, J., Gardina, S., Glade, S., Goyette, M., Graham, L.,
              Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
              Klein, J., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J.,
              Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
              McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McPheters, R.,
              McElrath, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
              Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
              O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
              Pisanli, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
              Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
              Stange-Thomann, N., Stojanovic, N., Sudranian, A., Talamas, J.,
              Testafey, S., Theodore, J., Thirell, A., Travers, M., Triggillo, J.,
              Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J.,
              Young, G., Zainoun, J., Zimmer, A. and Zody, M.
              Direct Submission
              Submitted (04-MAR-2000) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Mar 26, 2000 this sequence version replaced g1:7158126.
              All repeats were identified using RepeatMasker:
              Smit, A.F.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information
Center Project name: L7716
Center clone name: 217_0_2

```

NOTE: This record contains 98 individual sequencing reads that have not been assembled into contigs. Runs of N are used to separate the reads and the order in which they appear is completely arbitrary. Low-pass sequence sampling is useful for identifying clones that may be gene-rich and allow overlap relationships among clones to be deduced. However, it should not be assumed that this clone will be sequenced to completion. In the event that the record is updated, the accession number will be preserved.

```

1 740: contig of 740 bp in length
741 840: gap of 100 bp
841 1543: contig of 703 bp in length
1544 1643: gap of 100 bp
1644 2350: contig of 707 bp in length
2351 2450: gap of 100 bp
2451 3158: contig of 708 bp in length
3159 3258: gap of 100 bp
3259 3978: contig of 720 bp in length
3979 4078: gap of 100 bp
4079 4776: contig of 698 bp in length
4777 4876: gap of 100 bp
4877 5593: contig of 717 bp in length
5594 5693: gap of 100 bp
5694 6412: contig of 719 bp in length
6413 6512: gap of 100 bp
6513 7238: contig of 726 bp in length
7239 7338: gap of 100 bp
7339 8055: contig of 717 bp in length
8056 8155: gap of 100 bp
8156 8870: contig of 715 bp in length
8871 8970: gap of 100 bp
8971 9691: contig of 721 bp in length
9692 9791: gap of 100 bp
9792 10493: contig of 702 bp in length
10494 10593: gap of 100 bp
10594 11326: contig of 733 bp in length
11327 11426: gap of 100 bp
11427 12137: contig of 711 bp in length
12138 12237: gap of 100 bp
12238 12950: contig of 713 bp in length
12951 13050: gap of 100 bp
13051 13751: contig of 701 bp in length
13752 13851: gap of 100 bp
13852 14552: contig of 701 bp in length
14553 14652: gap of 100 bp
14653 15372: contig of 720 bp in length
15373 15472: gap of 100 bp
15473 16183: contig of 711 bp in length
16184 16283: gap of 100 bp
16284 17005: contig of 722 bp in length
17006 17105: gap of 100 bp
17106 17825: contig of 720 bp in length
17826 17925: gap of 100 bp
17926 18639: contig of 714 bp in length
18640 18739: gap of 100 bp
18740 19457: contig of 718 bp in length
19458 19557: gap of 100 bp
19558 20269: contig of 712 bp in length
20270 20369: gap of 100 bp
20370 21069: contig of 700 bp in length
21070 21169: gap of 100 bp
21170 21895: contig of 726 bp in length
21896 21995: gap of 100 bp
21996 22709: contig of 714 bp in length
22710 22809: gap of 100 bp
22810 23507: contig of 658 bp in length
23508 23607: gap of 100 bp

```

```

* 23608 24297: contig of 690 bp in length
* 24298 24397: gap of 100 bp
* 24398 25107: contig of 710 bp in length
* 25108 25207: gap of 100 bp
* 25208 25936: contig of 729 bp in length
* 25937 26036: gap of 100 bp
* 26037 26749: contig of 713 bp in length
* 26750 26849: gap of 100 bp
* 26850 27562: contig of 713 bp in length
* 27563 27662: gap of 100 bp
* 27663 28369: contig of 707 bp in length
* 28370 28469: gap of 100 bp
* 28470 29153: contig of 684 bp in length
* 29154 29253: gap of 100 bp
* 29254 29971: contig of 718 bp in length
* 29972 30071: gap of 100 bp
* 30072 30770: contig of 699 bp in length
* 30771 30870: gap of 100 bp
* 30871 31581: contig of 711 bp in length
* 31582 31681: gap of 100 bp
* 31682 32380: contig of 699 bp in length
* 32381 32480: gap of 100 bp
* 32481 33187: contig of 707 bp in length
* 33188 33287: gap of 100 bp
* 33288 34012: contig of 725 bp in length
* 34013 34112: gap of 100 bp
* 34113 34831: contig of 719 bp in length
* 34832 34931: gap of 100 bp
* 34932 35627: contig of 696 bp in length
* 35628 35727: gap of 100 bp
* 35728 36446: contig of 719 bp in length
* 36447 36546: gap of 100 bp
* 36547 37245: contig of 699 bp in length
* 37246 37345: gap of 100 bp
* 37346 38077: contig of 732 bp in length
* 38078 38177: gap of 100 bp
* 38178 38879: contig of 702 bp in length
* 38880 38979: gap of 100 bp
* 38980 39688: contig of 709 bp in length
* 39689 39788: gap of 100 bp
* 39789 40500: contig of 712 bp in length
* 40501 40600: gap of 100 bp
* 40601 41317: contig of 717 bp in length
* 41318 41417: gap of 100 bp
* 41418 42116: contig of 699 bp in length
* 42117 42216: gap of 100 bp
* 42217 42923: contig of 707 bp in length
* 42924 43023: gap of 100 bp
* 43024 43739: contig of 716 bp in length
* 43740 43839: gap of 100 bp
* 43840 44570: contig of 731 bp in length
* 44571 44670: gap of 100 bp
* 44671 45381: contig of 711 bp in length
* 45382 45481: gap of 100 bp
* 45482 46191: contig of 710 bp in length
* 46192 46291: gap of 100 bp
* 46292 46986: contig of 695 bp in length
* 46987 47086: gap of 100 bp
* 47087 47789: contig of 703 bp in length
* 47790 47889: gap of 100 bp
* 47890 48590: contig of 701 bp in length
* 48591 48690: gap of 100 bp
* 48691 49414: contig of 724 bp in length
* 49415 49514: gap of 100 bp
* 49515 50220: contig of 706 bp in length
* 50221 50320: gap of 100 bp
* 50321 51010: contig of 690 bp in length
* 51011 51110: gap of 100 bp
* 51111 51823: contig of 713 bp in length
* 51824 51923: gap of 100 bp
* 51924 52648: contig of 725 bp in length
* 52649 52748: gap of 100 bp
* 52749 53466: contig of 718 bp in length

```

---

```

Query Match      5.9%; Score 120; DB 2; Length 79785;
Best Local Similarity 100.0%; Pred. No. 9,1e-54;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

* 53467 53566: gap of 100 bp
* 53567 54295: contig of 729 bp in length
* 54296 54395: gap of 100 bp
* 54396 55105: contig of 710 bp in length
* 55106 55205: gap of 100 bp
* 55206 55935: contig of 730 bp in length

Db 25581 GAAGTATACACGTGTGACCTCTGTGATGCATTCATGCGATGGGAC 25522
Qy 1107 GAAAGTTATCACCTGTGTGACCTCTGTGATGCATTCATGCGATGGGAC 1166
Qy 1107 GAAAGTTATCACCTGTGTGACCTCTGTGATGCATTCATGCGATGGGAC 1166
Db 25581 GAAGTATACACGTGTGACCTCTGTGATGCATTCATGCGATGGGAC 25522

RESULT 15
BC019812
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gqabbs-remail.nih.gov
Tissue Procurement: jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amegbcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 41 Row: 1 Column: 20.
Location/Qualifiers
1. 2045
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:30541 IMAGE:5042856"
/tissue-type="Kidney, normal. 5 month old male mouse."
/clone_lib="NCI-CGAP_K1d14"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
406. 1386
/codon_start=1
/product="RIKEN CDNA 2310075G14 gene"
/protein_id="AAH19812.1"
/db_xref="GI:18044186"
/db_xref="LocusID:66966"

```

```

/translation="MGKGVVADKKVELEKEDGHELKKRSQVPEMAKALHPDKKRY
ARSLQVFEETGIESSEFLRQNAHEEGSGYQHGIFPNCIILMLADQALDELRDK
VDDMLAGLELEELRGFHRRIINLKNISNSODYQHGIFQSIGKFEFHEYLTBEKCIPE
TSQOLKLGIEALQVTKYRARKQNVKARKFLSPRSQDPAVYGLVSQSWKEES
LEPALINVOSEFIGQHKPTAMPVMAVMESNKSPYHNCDCDDI1IGDRPMWAKLSE
SHLHOLKRRRLDIDAVSATSQNSNSDCDPERIEGSSQHNQELKASY"

```

BASE COUNT  
ORIGIN

570 2

453 C

541 g

481 t

ORIGIN

Query Match  
Post 100%

2.4

SCO  
PR48;  
NO

10;  
30-1

Best Local Similarity  
Matches 48; Conser.

ive  
100-

Pr  
0;

No. \_\_\_\_\_  
matc \_\_\_\_\_

3e-1

Ind

Cap

0.

QY

365

GGGA  
|||||

ACCA

D'b

337

GGGA

**ACCPR**

Search completed: April 21, 2003, 21:27:10  
Job time : 6769 secs

Job time : 6769 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:23:36 ; Search time 3289 seconds

(Without alignments)  
10050.164 Million cell updates/sec

Title: US-09-513-151-3

Sequence: 1 CTCGCATAGATGCGCTCCG.....TTTACAAAGAAAAAAA 2041

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues.

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

EST: \*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	573	28.1	725	9	AI133396 HA2003 Hu
2	569	27.9	743	14	BM975520 UI-CF-EN1
3	541	26.5	578	14	BM847727 K-EST0127
4	529	25.9	725	10	BE315223 601141778
5	480	23.5	601	9	AI207688 HA3066 Hu
6	466	22.8	799	13	BI758117 603023811

7	451	22.1	1088	14	BM800217
8	433	21.2	869	12	BG612651
9	422	20.7	639	10	AV759288
10	388	19.0	559	9	AI1718450
11	387	19.0	624	14	B0003256
12	387	19.0	783	13	BI222788
13	385	18.9	457	9	AA121465
14	372	18.2	543	9	AI193562
15	371	18.2	562	12	BG495857
16	371	18.2	963	13	BI462398
17	370	18.1	490	9	AI127851
18	367	18.0	487	9	AI337829
19	367	18.0	724	14	BM721352
20	364	17.8	758	13	BI820461
21	361	17.7	469	12	BF376938
22	357	17.5	528	12	BF064275
23	352	17.2	430	9	AI089239
24	350	17.1	350	9	AA847885
25	350	17.1	465	9	AA767411
26	342	16.8	450	9	AA888867
27	342	16.8	513	10	AW104415
28	340	16.7	440	14	N26804
29	340	16.7	444	14	N26803
30	340	16.7	473	14	N29610
31	339	16.6	442	9	AA830847
32	338	16.6	413	10	AW193859
33	338	16.6	443	14	N41610
34	337	16.5	386	10	AM197736
35	337	16.5	446	9	AA121308
36	335	16.4	433	9	AA716699
37	329	16.1	500	9	AI192804
38	323	15.8	711	12	BG506209
39	321	15.7	643	10	AW953634
40	320	15.7	554	9	AI749267
41	314	15.4	772	12	BG034660
42	309	15.1	836	12	BG575745
43	305	14.9	922	9	AL530575
44	301	14.7	434	9	AI347522
45	301	14.7	593	10	AW444470

## ALIGNMENTS

RESULT 1  
LOCUS AI133396/c 725 bp mRNA linear EST 11-NOV-1999  
DEFINITION HA2003 Human fetal liver CDNA library Homo sapiens CDNA, mRNA  
ACCESSION AI133396  
VERSION AI133396.1 GI:6360712  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M. and He, F.  
Expression profile analysis of a human fetal liver CDNA library  
Unpublished (1998)  
Contact: Yongtao Yu  
Department of Hematology  
Beijing Institute of Radiation Medicine  
27 Taiping Road, Beijing 100850, P.R.China  
Tel: 0086-10-68159479  
Fax: 0086-10-68214653  
Email: yyt48@yahoo.com.  
FEATURES  
source  
1..725  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="Human fetal liver CDNA library"

```

/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/p3"
/notes="Vector: pCDNA1"
BASE COUNT      220 a      177 c      124 g      204 t
ORIGIN

Query Match      28.1%; Score 573; DB 9; Length 725;
Best Local Similarity 99.9%; Pred. No. 2.8e-202;
Matches 693; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1296 CCCCCGAGCAATGATCAAGAGCTGAATGACGCGTTTAAGACATGTCAGTGCGCT 1355
      |||
DB 694 CCCCCGAGCAATGATCAAGAGCTGAATGACGCGTTTAAGACATGTCAGTGCGCT 635
      |||
OY 1356 TTGGAAGGCTGGGGGATCCCTTCAGAGGAGGGGATGTTTGTCTCCAGTGGG 1415
      |||
DB 634 TTGGAAGGCTGGGGGATCCCTTCAGAGGAGGGGATGTTTGTCTCCAGTGGG 575
      |||
OY 1416 CAAAGAGTCTTATGCGGATTCCTGATAGCAGAAACCTCCACATTTCTTTGA 1475
      |||
DB 574 CAAAGAGTCTTATGCGGATTCCTGATAGCAGAAACCTCCACATTTCTTTGA 515
      |||
OY 1476 TGTGCTTTTAAAGTCTCAGCTTCTATATAATAGAAACAGCAGGCTCTTGCCTCTGT 1535
      |||
DB 514 TGTGCTTTTAAAGTCTCAGCTTCTATATAATAGAAACAGCAGGCTCTTGCCTCTGT 455
      |||
OY 1536 GTGCGTGAATGTCGGAATGATGATGTCAGAGAAAGCATTTTCTTTTGAACCT 1595
      |||
DB 454 GTGCGTGAATGTCGGAATGATGATGTCAGAGAAAGCATTTTCTTTTGAACCT 395
      |||
OY 1596 TAAAGCTTATTTTAAAGCAGCAGATTCACATTTTATACATGAGATCTTCT 1655
      |||
DB 394 TAAAGCTTATTTTAAAGCAGCAGATTCACATTTTATACATGAGATCTTCT 335
      |||
OY 1656 TGTGCTGAATGTCGGAATGATGATGTCAGAGAAAGCATTTTCTTTTGAACCT 1715
      |||
DB 334 TGTGCTGAATGTCGGAATGATGATGTCAGAGAAAGCATTTTCTTTTGAACCT 276
      |||
OY 1716 GCTAATATATCTAATTTTCAGATGCTTTTGTAGATGATGATGATGATGATGAT 1775
      |||
DB 275 GCTAATATATCTAATTTTCAGATGCTTTTGTAGATGATGATGATGATGATGAT 216
      |||
OY 1776 ATTTGGAGTTCTAGATTTTGTAGATGATGATGATGATGATGATGATGATGAT 1835
      |||
DB 215 ATTTGGAGTTCTAGATTTTGTAGATGATGATGATGATGATGATGATGATGAT 156
      |||
OY 1836 GTGACCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1895
      |||
DB 155 GTGACCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 96
      |||
OY 1896 TGACATAGACTTGAAGACCAAGATTTGCAAGTCTGCAATGCTGCAATGCTG 1955
      |||
DB 95 TGACATAGACTTGAAGACCAAGATTTGCAAGTCTGCAATGCTGCAATGCTG 36
      |||
OY 1956 TATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1989
      |||
DB 35 TATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2
      |||
RESULT 2
BM975520/c 743 bp mRNA linear EST 21-MAR-2002
LOCUS      BM975520
DEFINITION UI-CF-EN1-acw-c-07-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
ACCESSION BM975520
VERSION   BM975520.1 GI:19593111
KEYWORDS  EST.
SOURCE    human.
ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 743)

```

```

AUTHORS      Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE        Normalization and subtraction: two approaches to facilitate gene
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
MEDLINE      97044477
COMMENT      Contact: McCray, PB
              University of Iowa
              2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
              Tel: 319 356 4866
              Fax: 319 356 7171
              Email: paul-mccray@uiowa.edu
              Tissue Procurement: Dr. M. J. Welsh, University of Iowa
              cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.resgen.com).
              The following repetitive elements were found in this cDNA
              sequence: 1-56, >NT_116110_low_complexity (matched complement)
              Seq primer: M13 FORWARD
              Polya=yes.

FEATURES
    source          Location/Qualifiers
    1..743
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="UI-CF-EN1-acw-c-07-0-UI"
    /clone_11b="UI-CF-EN1"
    /tissue_type="Primary Lung Cystic Fibrosis Epithelial
    cells"
    /dev_stage="Adult"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /note="Organ: Lung; Vector: pPVT3-Pac (Pharmacia) with a
    modified polylinker; Site_1: EcoR I; Site_2: Not I;
    UI-CF-EN1 is a normalized cDNA library containing the
    following tissue(s): Primary Lung Cystic Fibrosis
    Epithelial Cells. The library was constructed according to
    Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
    1996. First strand cDNA synthesis was primed with an
    oligo-dT primer containing a Not I site. Double stranded
    cDNA was ligated to an EcoR I adaptor, digested with Not
    I, and cloned directionally into pPVT3-Pac vector. The
    oligonucleotide used to prime the synthesis of
    first-strand cDNA contains a library tag sequence that is
    located between the Not I site and the (dT)18 tail. The
    sequence tag for this library is CTGCTCAGT.
    TAG_L1B=UI-CF-EN1
    TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
    6hr to LPS 24h
    TAG_SEQ=CTGCTCAGT"

BASE COUNT      231 a      172 c      116 g      224 t
ORIGIN

Query Match      27.9%; Score 569; DB 14; Length 743;
Best Local Similarity 99.7%; Pred. No. 8.3e-201;
Matches 739; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 1301 GGGCAATGATCAAGAGCTGAATGACGCGTTTAAGACATGTCAGTGCGCTTTGA 1360
      |||
DB 743 GGGCAATGATCAAGAGCTGAATGACGCGTTTAAGACATGTCAGTGCGCTTTGA 684
      |||
OY 1361 AAGGTGCTGGGATCCATTCAGAGGAGGAGGATGTTTGTCTCCAGTCTGGGCAAG 1420
      |||
DB 683 AAGGTGCTGGGATCCATTCAGAGGAGGAGGATGTTTGTCTCCAGTCTGGGCAAG 624
      |||
OY 1421 GAGTCTATGCGGAATTCCTGATAGCAGAAAGCTCCACATTTCTTTGATGTGG 1480
      |||
DB 623 GAGTCTATGCGGAATTCCTGATAGCAGAAAGCTCCACATTTCTTTGATGTGG 564
      |||
OY 1481 TTTTAAAGTCTCAGCTTCTATATAATAGAAACAGCAGGCTTGTGACGCTCTGTGTGGC 1540
      |||
DB 563 TTTTAAAGTCTCAGCTTCTATATAATAGAAACAGCAGGCTTGTGACGCTCTGTGTGGC 504
      |||

```





ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE NIH-MGC http://mgc.ncl.nih.gov/  
 1 (bases 1 to 725)  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgsbds-remail.nih.gov  
 Tissue Procurement: DCTD/DRP  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov  
 Plate: LCM111 row\* k column: 20  
 High quality sequence stop: 720.

FEATURES  
 Location/Qualifiers  
 1..725  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3141643"  
 /clone\_lib="NIH-MGC\_9"  
 /tissue\_type="adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2:  
 EcoRI; CDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGGCAGG(G). Size-selected >500bp for average  
 insert size 1.8kb. Library constructed by Ling Hong in  
 the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-CDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."  
 BASE COUNT 186 a 175 c 197 g 167 t  
 ORIGIN

Query Match 25.9%; Score 529; DB 10; Length 725;  
 Best Local Similarity 99.7%; Pred. No. 5.6e-186;  
 Matches 699; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

28 TGACGAGCAGTTCCTGCGGAGTGGGCTCAGGGGCTGCACAGCAGCCTTCTTGT 87  
 1 TGCACGACGACTTCTCTGCGAGTGGGCTCAGGGGCTGCACAGCAGCCTTCTTGT 60  
 88 ACTGATTTCTCGGGCCACGCGGCAATCCAGCCTGGCTGAGCTAGGCCAGCG 147  
 61 AGTATTTCTCGGGCCACGCGGCAATCCAGCCTGGCTGAGCTAGGCCAGCG 120  
 148 GCTCGGCGGTGAGATGTCAGCGCTGATCCATGCAAGCTTATGAAGGCTTAGACATCAT 207  
 121 GCTCGGCGGTGAGATGTCAGCGCTGATCCATGCAAGCTTATGAAGGCTTAGACATCAT 180  
 208 CACCAACAAGTTTCTGCCAGAGAGATCTCCCGACACCATGATGACGTTTGT 267  
 181 CACCAACAAGTTTCTGCCAGAGAGATCTCCCGACACCATGATGACGTTTGT 229  
 268 GATCCTCTTGAGCAATTTACACAGTGGGCTTCAAAATAGAGCAACTCTGTAT 327  
 240 GATCCTCTTGAGCAATTTACACAGTGGGCTTCAAAATAGAGCAACTCTGTAT 299  
 328 TGAATATATTTTGGCCGAGCAAAATCTTATTTGTTGGGAGAAACATTTATACAT 387  
 300 TGAATATATTTTGGCCGAGCAAAATCTTATTTGTTGGGAGAAACATTTATACAT 359  
 388 TGAATATCTGCTCTGGAAGATTCTTGTCAATACCAAGCCCAAGAGATGGGACTGAGA 447  
 360 TGAATATCTGCTCTGGAAGATTCTTGTCAATACCAAGCCCAAGAGATGGGACTGAGA 419  
 448 AGTATGACGAGAAAGTGGAGTGAAGAAGAGATGGTCTTGTACTTCAAAAGCGCT 507  
 420 AGTATGACGAGAAAGTGGAGTGAAGAAGAGATGGTCTTGTACTTCAAAAGCGCT 479  
 508 AAGCAGGTGACCAAGAAATGCTGCCAAGCTGCATCCATGACAAACGAAAGTGC 567

Db 480 AAGCAGGTGACCAAGAAATGCTGCCAAGCTGCATCCATGACAAACCAAGTGC 539  
 568 CAGAGCTTGCAGATTTTGAAGAAGAGAAATCTCATAGTAATTTCTCATGCTCA 627  
 540 CAGAGCTTGCAGATTTTGAAGAAGAGAAATCTCATAGTAATTTCTCATGCTCA 599  
 628 ACATACGAGAGAGAGTGGTCCCTTGCAGAGTCTCTGAAGTTCTCTAACCTTCAT 687  
 600 ACATACGAGAGAGAGTGGTCCCTTGCAGAGTCTCTGAAGTTCTCTAACCTTCAT 659  
 688 CCTTGGCTTCATGCTGACGAGCAGGATTTAGATGAGCGCT 728  
 660 CCTTGGCTTCATGCTGACGAGCAGGATTTAGATGAGCGCT 700

RESULT 5  
 AI207688/c 601 bp mRNA linear EST 11-NOV-1999  
 LOCUS  
 DEFINITION HA3066 Human fetal liver cDNA library Homo sapiens CDNA, mRNA  
 sequence.  
 ACCESSION AI207688  
 VERSION AI207688.1 GI:6361702  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.  
 and He, F.  
 Expression profile analysis of a human fetal liver cDNA library  
 Unpublished (1998)  
 CONTACT: Chenggang Zhang  
 Beijing Institute of Radiation Medicine  
 27 Taiping Road, Beijing 100850, P.R. China  
 Email: zhang\_chenggang@hotmail.com.

FEATURES  
 Location/Qualifiers  
 1..601  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Human fetal liver cDNA library"  
 /tissue\_type="liver"  
 /dev\_stage="fetal"  
 /lab\_host="MC1061/P3"  
 /note="Vector: pCDNA1"  
 BASE COUNT 198 a 137 c 102 g 164 t  
 ORIGIN

Query Match 23.5%; Score 480; DB 9; Length 601;  
 Best Local Similarity 99.8%; Pred. No. 9.5e-168;  
 Matches 600; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

1389 AAGGATATGTTTGTCTCCAGCTGCGGCAAGAGAGTATCGCGAATCTCTGCATAC 1448  
 601 AAGGATATGTTTGTCTCCAGCTGCGGCAAGAGAGTATCGCGAATCTCTGCATAC 542  
 1449 AGAAAGCTCCACCATTTTCTTTGATGTGTTTAAAGTCTCAAGCTCTCTAATAAG 1508  
 541 AGAAAGCTCCACCATTTTCTTTGATGTGTTTAAAGTCTCAAGCTCTCTAATAAG 482  
 1509 AAACAGCAGCTTGTGACCTCTTGTGCTGATGCTGGAAGATGATGATTCAG 1568  
 481 AAACAGCAGCTTGTGACCTCTTGTGCTGATGCTGGAAGATGATGATTCAG 422  
 1569 GAAAGCATTTTCTTTGTAACCTTAAGCTCTATTAATAAGCAGACAGATTG 1628  
 421 GAAAGCATTTTCTTTGTAACCTTAAGCTCTATTAATAAGCAGACAGATTG 362  
 1629 CACATTTTATACATGAGATCTTCTTGTGCTGAATACAGATTTGATGATCCCTTT 1688  
 361 CACATTTTATACATGAGATCTTCTTGTGCTGAATACAGATTTGATGATCCCTTT 302

QY 1689 AAAAGAGTTTATGTCCTGAGCTGGCTAAATATCTAATTTCCAGATGCTTTGTA 1748  
 |||||||  
 Db 301 -AAAGAGTTTATGTCCTGAGCTGGCTAAATATCTAATTTCCAGATGCTTTGTA 243  
 |||||||  
 QY 1749 GATGACTGAAGTATTTTGAGCCACATATTGGAGTTTGAATTTGATGAATGGCAGGA 1808  
 |||||||  
 Db 242 GATGACTGAAGTATTTTGAGCCACATATTGGAGTTTGAATTTGATGAATGGCAGGA 183  
 |||||||  
 QY 1809 AAGGCCATCTCCATTGATGATTAAAGTAAACCAACTAGTTCTCGGAATTTCTACAGAG 1868  
 |||||||  
 Db 182 AAGGCCATCTCCATTGATGATTAAAGTAAACCAACTAGTTCTCGGAATTTCTACAGAG 123  
 |||||||  
 QY 1869 AAGGCCATCTCCATTGATGATTAAAGTAAACCAACTAGTTCTCGGAATTTCTACAGAG 1928  
 |||||||  
 Db 122 AAGGCCATCTCCATTGATGATTAAAGTAAACCAACTAGTTCTCGGAATTTCTACAGAG 63  
 |||||||  
 QY 1929 TTTCGAGCTGCTCATGTGAGTATTATTCACGTGCTTTCTATTGATGATCAATC 1988  
 |||||||  
 Db 62 TTTCGAGCTGCTCATGTGAGTATTATTCACGTGCTTTCTATTGATGATCAATC 3  
 |||||||  
 QY 1989 T 1989  
 |||||||  
 Db 2 T 2

RESULT 6  
 BI758117 799 bp mRNA linear EST 25-SEP-2001  
 LOCUS 60302381F1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5194347 5'  
 DEFINITION  
 BI758117  
 mRNA sequence.  
 BI758117  
 EST.  
 BI758117.1 GI:15749695  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 799)

REFERENCE  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rgs@bbs-rtmail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov

plate: LAM1486 row: e column: 04  
 High quality sequence start: 2  
 High quality sequence stop: 775.  
 Location/Qualifiers

## FEATURES

1..799  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5194347"  
 /clone\_id="NIH\_MGC\_114"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6  
 male brains, age range 23-27 yo. Library is oligo-dT  
 primed and directionally cloned (EcoRV site is destroyed  
 upon cloning). Average insert size 1.5 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 019. Note:  
 this is a NIH MGC Library."

BASE COUNT 207 a 188 c 220 g 184 t  
 ORIGIN

Query Match 22.8%; Score 466; DB 13; Length 799;

Best Local Similarity 99.6%; Pred. No. 1,1e-162;  
 Matches 756; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 12 TGGGTCGCGGGGGGCGACAGACAGTTCCTGGGACATGGAGCTCAGGGCCCTGCAAC 71  
 |||||||  
 Db 1 TGGCGTCGCGGGGGGCGACAGACAGTTCCTGGGACATGGAGCTCAGGGCCCTGCAAC 60  
 |||||||  
 QY 72 GGAACCTACCTCTTGTAGTATTCGCGGCGACAGGGCCACGGCAAAATCCAGCTGGCGT 131  
 |||||||  
 Db 61 GGAACCTACCTCTTGTAGTATTCGCGGCGACAGGGCCACGGCAAAATCCAGCTGGCGT 120  
 |||||||  
 QY 132 TGCAGTAGGCGACAGGCGCTCGGCGGTAGATCGTACGCTGACTCCATGACGTCTATG 191  
 |||||||  
 Db 121 TGCAGTAGGCGACAGGCGCTCGGCGGTAGATCGTACGCTGACTCCATGACGTCTATG 180  
 |||||||  
 QY 192 AAGGCTTAGACATCATACCAACAAGTTTCTGCGCAAGACAGAGATCGCGCGACCC 251  
 |||||||  
 Db 181 AAGGCTTAGACATCATACCAACAAGTTTCTGCGCAAGACAGAGATCGCGCGACCC 240  
 |||||||  
 QY 252 ACATGATCAGCTTGTGATCCTCTTGTGACCAATTTACAGAGTGTGAGCTTGAGAAATA 311  
 |||||||  
 Db 241 ACATGATCAGCTTGTGATCCTCTTGTGACCAATTTACAGAGTGTGAGCTTGAGAAATA 300  
 |||||||  
 QY 312 GAGCACTGCTGTGATTGAGATATATTGGCCGAGACAAATTTCTATTGTTGGGAG 371  
 |||||||  
 Db 301 GAGCACTGCTGTGATTGAGATATATTGGCCGAGACAAATTTCTA-75TGTGGAG 359  
 |||||||  
 QY 372 GAACCAATTTATGATGAAATCTGCTGGGAAGTTCTGTCAATACCAAGCCCGAG 431  
 |||||||  
 Db 360 GAACCAATTTATGATGAAATCTGCTGGGAAGTTCTGTCAATACCAAGCCCGAG 419  
 |||||||  
 QY 432 AGATGGGACATGAGAAAGTGAATGACGAAAGTGAAGCTTGAAAGAGAGATGGTCTG 491  
 |||||||  
 Db 420 AGATGGGACATGAGAAAGTGAATGACGAAAGTGAAGCTTGAAAGAGAGATGGTCTG 479  
 |||||||  
 QY 492 TACTTCACAAAGCCTTAAGCCAGGTGACCCAGAAATGCTGCCAAGCTGATCCATG 551  
 |||||||  
 Db 480 TACTTCAC- AAGCCTTAAGCCAGGTGACCCAGAAATGCTGCCAAGCTGATCCATG 538  
 |||||||  
 QY 552 ACAAAAGCAAGTGGGCGGAGAGCTTGCAAGTTTGAAGAAACAGAAATCTCTATAGT 611  
 |||||||  
 Db 539 ACAAAAGCAAGTGGGCGGAGAGCTTGCAAGTTTGAAGAAACAGAAATCTCTATAGT 598  
 |||||||  
 QY 612 AATTTCATCATGTCACATACGGAAGAGTGTGCTGCCCTTGAGAGTCTCTGAAT 671  
 |||||||  
 Db 599 AATTACATCATGTCACATACGGAAGAGTGTGCTGCCCTTGAGAGTCTCTGAAT 658  
 |||||||  
 QY 672 TCTCTAACCTTGCATCTCTTGGCTTATGCTGACACGAGCTTACATAGGCTTGG 731  
 |||||||  
 Db 659 TCTCTAACCTTGCATCTCTTGGCTTATGCTGACACGAGCTTCTAGGTAGCGCTTGG 718  
 |||||||  
 QY 732 ATAAGAGGTTGATGACATGCTGCTGCTGGGCTCTTG 770  
 |||||||  
 Db 719 ATAAGAGGTTGATGACATGCTGCTGCTGGGCTCTTG 757  
 |||||||

RESULT 7  
 BM800217 1088 bp mRNA linear EST 05-MAR-2002  
 LOCUS AGNCOURT\_6416157 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:5531248  
 DEFINITION  
 AGNCOURT\_6416157 NIH\_MGC\_71  
 5', mRNA sequence.  
 BM800217  
 EST.  
 BM800217.1 GI:19117040  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 1088)

REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT

Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

Plate: LLM12212 row: j column: 17  
High quality sequence stop: 651.

## FEATURES

source

Location/Qualifiers  
1. 1088

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5531248"

/clone\_lib="NIH\_MGC\_71"

/tissue\_type="leiomyosarcoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; NotI;  
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."

BASE COUNT 318 a 218 c 255 g 294 t 3 others

## ORIGIN

Query Match

Best Local Similarity 99.4%; Score 451; DB 14; Length 1088;

Matches 651; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

## FEATURES

source

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4771338"

/clone\_lib="NIH\_MGC\_61"

/tissue\_type="embryonal carcinoma"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1;  
SfiI (ggccgagctggcc); Site: 2; SfiI (ggccatcagctg);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-ATCTCAGGCGCGCCGCGCATG-3' and 3'  
sequence: 5'-ATCTCAGGCGCGCCGCGCATG-3' (30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library."

BASE COUNT 259 a 184 c 206 g 220 t

## ORIGIN

Query Match

Best Local Similarity 99.6%; Score 433; DB 12; Length 869;

Matches 533; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 8  
BG612651

LOCUS BG612651 869 bp mRNA linear EST 18-APR-2001

DEFINITION 602640078F1 NIH\_MGC\_61 Homo sapiens cDNA clone IMAGE:4771338 5',  
mRNA sequence.

ACCESSION BG612651

VERSION BG612651.1 GI:13664022

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 869)  
NIH\_MGC http://mgi.ncl.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: CLONETECH Laboratories, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM1639 row: k column: 19

High quality sequence stop: 684.

Location/Qualifiers  
1. 869

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4771338"

/clone\_lib="NIH\_MGC\_61"

/tissue\_type="embryonal carcinoma"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site: 1;  
SfiI (ggccgagctggcc); Site: 2; SfiI (ggccatcagctg);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-ATCTCAGGCGCGCCGCGCATG-3' and 3'  
sequence: 5'-ATCTCAGGCGCGCCGCGCATG-3' (30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library."

BASE COUNT 259 a 184 c 206 g 220 t

## ORIGIN

Query Match

Best Local Similarity 99.6%; Score 433; DB 12; Length 869;

Matches 533; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1048 GGGCCACAGGCTTACAGCCACTGCATTAAGATGCCATTCATTAAGCTGACACACAG 1107

1 GGGCCACAGGCTTACAGCCACTGCATTAAGATGCCATTCATTAAGCTGACACACAG 60

1108 AAGTATACCTGTGTGACCTGTGTGATGATCATTCATTTGGGATTCGGGATTCGG 1167

61 AAGTATACCTGTGTGACCTGTGTGATGATCATTCATTTGGGATTCGGGATTCGG 120

1168 GCACATTAATCCCAATCCCACTTGAACCACTGAAGAAAGAAAGAAAGATTCGACTG 1227

121 GCACATTAATCCCAATCCCACTTGAACCACTGAAGAAAGAAAGATTCGACTG 180

1228 TGCCTCAACACATTAAGAAAGTCAAGTCTTCCCACTATTAAGAAAGTCAAG 1287

181 TGCCTCAACACATTAAGAAAGTCAAGTCTTCCCACTATTAAGAAAGTCAAG 240

1288 GAAGGATCCCGAGGCGCAATGATCAAGAGCTGAATGACGCTTAAAGACATGTC 1347

241 GAAGGATCCCGAGGCGCAATGATCAAGAGCTGAATGACGCTTAAAGACATGTC 300

```

/Organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1IMAGE="2332692"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/note="Organ: Colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR1; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGATCTGAAGTCGAGCGCCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCACTACTAAT 3' and 5' ATTCACTGTG 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead."
BASE COUNT      191 a      118 C      84 G      166 t
ORIGIN
Query Match      19.0%; Score 388; DB 9; Length 559;
Best Local Similarity 99.6%; Pred. No. 1.2e-133;

```

	Matches	558:	Conservative	0:	Mismatches	1:	Indels	1:	Gaps	1:
QY	1470	TTTTATGTGGTTTAAAGTCTCACGTTCTATATACAAACAGCAGCTCTTGCAGCT								1529
Db	559	TTTTATGTGGTTTAAAGTCTCACGTTCTCTATATAGAAACACAGCTCTTGCAGCT								500
QY	1530	CCTTGTGGCGTGAATGTGCTGGAAATAGTATGATTCAGGAAGCATTTTTTTTCTTT								1589
Db	499	CCTTGTGGCGTGAATGTGCTGGAAATAGTATGATTCAGGAAGCATTTTTTTTCTTT								440
QY	1590	GAACCTTAAAGGTTCTATTATTTAAAGCAGCAGACATTCACATTTTATACATGAGAT								1649
Db	439	GAACCTTAAAGGTTCTATTATTTAAAGCAGCAGACATTCACATTTTATACATGAGAT								380
QY	1650	CTTCTTTGTGGTGAATACCAGAGTTGACTGCATCCCTTTAAAGAAGTTTATGTCCCTG								1709
Db	379	CTTCTTTGTGGTGAATACCAGAGTTGACTGCATCCCTTT - AAAGAAGTTTATGTCCCTG								321
QY	1710	ACTCGGCTAAAAATATCTAATTTCCAGATGGTTTGTAGAGACTGAAGTATTTTCTGAG								1769
Db	320	ACTCGGCTAAAAATATCTAATTTCCAGATGGTTTGTAGAGACTGAAGTATTTTGTGAG								261
QY	1770	CCACATATTGGGAGTTCTAGATTGAGTGAATGGCAGAAAGGCCATCTCCATTGAGAT								1829
Db	260	CCACATATTGGGAGTTCTAGATTGAGTGAATGGCAGAAAGGCCATCTCCATTGAGAT								201
QY	1830	GATTTAAGTGAACCAACTAGTCTCTCGGAATCTACAGAGAAAGAGGAAATCAGACTGAGC								1889
Db	200	GATTTAAGTGAACCAAACTAGTCTCTCGGAATCTACAGAGAAAGAGGAAATCAGACTGAGG								141
QY	1890	AAGCTGTACATAGACTTTGAAGACCAAAAGCTTTGAAATTTGGCAGCGTCATAGTGTG								1949
Db	140	AAGCTGTACATAGACTTTGAAGACCAAAAGCTTTGAAATTTGGCAGCGTCATAGTGTG								81
QY	1950	AGTATTATCACTGCTGTCTTCTATTAGTGTACAAATCTATATTTTATTCGAAGTTTAA								2009
Db	80	AGTATTATCACTGCTGTCTTCTATTAGTGTACAAATCTATATTTTATTCGAAGTTTAA								21
QY	2010	ATAAAGAAAAAATTTCAG 2029								
Db	20	ATAAAGAAAAAATTTCAG 1								

RESULT	11
BQ003256/c	
LOCUS	BQ003256
DEFINITION	624 bp mRNA linear EST 26-MAR-2002 U1-H-E11.ayx-n-12-0-U1.s1 NC1 CGAP_E11 Homo sapiens CDNA clone IMAGE:5845067 3', mRNA sequence.
ACCESSION	BQ003256

ACCESSION	BQ003250
VERSION	BQ003256.1
KEYWORDS	GI:19728156
SOURCE	EST.
ORGANISM	human.
	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 624)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, ph.D.

Email: cgapds-remail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
The following repetitive elements were found in this cDNA sequence: 1-36, >POLY\_A#Simple\_repeat (matched complement)  
Seq primer: M13 FORWARD  
POLYA=yes.

```

source
1..624
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5845067"
/clone_lib="NCI_CGAP_E11"
/tissue_type="Chondrosarcoma"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/notes="Organ: Left Pelvis; Vector: pPTT3-Pac (Pharmacia)
with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI_CGAP_E11 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma. The library was
constructed according to Bonaligo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
ACACTGCAC.
TAG_LIB-UT-H-E11
TAG_TISSUE=chondrosarcoma
TAG_SEQ=ACACTGCAC"
BASE COUNT      203 a      128 c      97 g      195 t      1 others
ORIGIN
19.0%: Score 387; DB 14; Length 624;
Query Match      99.5%; Pred. 2.6e-13;
Matches 607; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```

OY	1419	AGGATGTCTATGCGGAATTCTCTGCATGACAGAAAAAGTCCACCATTCTTTTGATGT	1478
Dd	624	AGGATGTCTATGCGGAATTCTCTGCATGACAGAAAAAGTCCACCATTCTTTTGATGT	565
OY	1479	GSTTTAAAGTCACAGTCTCTATAAATGAAGAAGCATGTTCTCAGCTCTGTGTG	1538
Dd	564	GSTTTAAAGTCACAGTCTCTATAAATGAAGAAGCATGTTCTCAGCTCTGTGTG	505
OY	1539	GCTGATGTCTGCAAAATGATGTAGTTCAGGAACAATTTTTTTTTCTTTGAACTTAA	1598
Dd	504	GCTGATGTCTGCAAAATGATGTAGTTCAGGAACAATTTTTTTTTCTTTGAACTTAA	445
OY	1599	AGGTTCTATTATTAAAGCACGACAGATTCCACATTTTAAATCATGAGATCTCTTGT	1658
Dd	444	AGGTTCTATTATTAAAGCACGACAGATTCCACATTTTAAATCATGAGATCTCTTGT	385
OY	1659	GCTGATATCCAGAGATGTACTGCATCCCTTTAAAACAATTTAATGCCCTGCTGGCT	1718
Dd	384	GCTGATATCCAGAGATGTACTGCATCCCTTTAAAACAATTTAATGCCCTGCTGGCT	326
OY	1719	AAAATTATCTAATTTCCAGATGCTTTTGTAGTAGCTAGTAATTTGTGAGCCACATATT	1778
Dd	325	AAAATTATCTAATTTCCAGATGCTTTTGTAGTAGCTAGTAATTTGTGAGCCACATATT	266
OY	1779	GGGAGTTCTAATTTGATGATGAATGGCAGAGAAAGGCCATTCCTATTGAGATGATTAAATG	1838
Dd	265	GGGAGTTCTAATTTGATGATGAATGGCAGAGAAAGGCCATTCCTATTGAGATGATTAAATG	206
OY	1839	AACCAAATCTAGTTCTCGGAATTTCTACAGAGAAGAGGGAATCCAGACTAGGAAGCTGTGA	1898
Dd	205	AACCAAATCTAGTTCTCGGAATTTCTACAGAGAAGAGGGAATCCAGACTAGGAAGCTGTGA	146
OY	1899	CATAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTATGTGTAGTTATTAT	1958
Dd	145	CATAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTATGTGTAGTTATTAT	86
OY	1959	CACGCTCTCTTTTCATGTAGAGTTACAATCTATATTTTATTTGAAGTTAAATTAAGAAA	2018
Dd	85	CACGCTCTCTTTTCATGTAGAGTTACAATCTATATTTTATTTGAAGTTAAATTAAGAAA	26
OY	2019	AAATTTACAA 2028	

```

Db      25 AAATTRCAA 16
|||||
RESULT 12
BI222788      783 bp  mRNA  linear  EST 11-JUL-2001
LOCUS        602941574F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5104483 5',
DEFINITION
ACCESSION    BI222788
VERSION      BI222788
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 783)
AUTHORS      NIH-MGC http://mgc.ncl.nih.gov/
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: c9apds-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LLM1252 row: d column: 20
              High quality sequence stop: 741.
FEATURES
  source
    1..783
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5104483"
    /clone_lib="NIH_MGC_12"
    /tissue_type="cervical carcinoma cell line"
    /lab_host="DH10B"
    /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
    Average insert size 1.4 kb. Library prepared by Life
    Technologies."
BASE COUNT   218 a 181 c 184 g 200 t
ORIGIN
Query Match      19.0%; Score 387; DB 13; Length 783;
Best Local Similarity 99.8%; Pred. No. 2.1e-133;
Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
OY 203 ATCATCCACCAACAGGTTTCCCAAGACAGAGATCTGCCGCCACCACATGATCAGC 262
    |||||||
Db 1 ATCATCCACCAACAGGTTTCCCAAGACAGAGATCTGCCGCCACCACATGATCAGC 60
OY 263 TTGTGGATCCTCTGTGACCAATACACAGTGTGGACTTGTAGAAATAGACCACTGCT 322
    |||||||
Db 61 TTGTGGATCCTCTGTGACCAATACACAGTGTGGACTTGTAGAAATAGACCACTGCT 120
OY 322 CGATTGAAGATATATTTGGCCGAGACAATAATCTATGTGTGGAGAGACCAATAT 382
    |||||||
Db 121 CTGATTGAAGATATATTTGGCCGAGACAATAATCTATGTGTGGAGAGACCAATAT 180
OY 383 TACATTGAATCTCTGCTGGAAGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACT 442
    |||||||
Db 181 TACATTGAATCTCTGCTGGAAGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACT 240
OY 443 GAGAAAGTATGACCGAAAGTGAAGCTTGA- AAGAGGATGCTCTTGTCTACAA 501
    |||||||
Db 241 GAGAAAGTATGACCGAAAGTGAAGCTTGA- AAGAGGATGCTCTTGTCTACAA 300
OY 502 AGCCCTAAGCCAGGTGGAGCCAGAAATGCGTCAAGTGCATCCATGACAAACGCAA 561
    |||||||
Db 301 AGCCCTAAGCCAGGTGGAGCCAGAAATGCGTCAAGTGCATCCATGACAAACGCAA 360

```

```

OY 562 AGTGCCAGAGAGCTTGCAAGTTTGTGAAGAAGAGAAATCTCATAGTAATTTCTCCA 621
    |||||||
Db 361 AGTGCCAGAGAGCTTGCAAGTTTGTGAAGAAGAGAAATCTCATAGTAATTTCTCCA 420
OY 622 TCGTCAACATACGAGAGAGAGTGTGTGCTTGGAGCTCTCTGTAAGTCTCTAACCC 661
    |||||||
Db 421 TCGTCAACATACGAGAGAGAGTGTGTGCTTGGAGCTCTCTGTAAGTCTCTAACCC 480
OY 682 TTGCATCCCTTTGGCTTCATGCTGACAG 709
    |||||||
Db 481 TTGCATCCCTTTGGCTTCATGCTGACAG 508
RESULT 13
AA121465      457 bp  mRNA  linear  EST 11-NOV-1996
LOCUS        2X91C07.r1 Soares-pregnant-uterus_NbH9U Homo sapiens cDNA clone
DEFINITION
ACCESSION    AA121465
VERSION      AA121465
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 457)
AUTHORS      Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
              'M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
              Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston
              'R., Williamson, A., Wohlmann, P. and Wilson, R.
              The WashU-Merck EST Project
              Unpublished (1995)
              Contact: Wilson, R.K.
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              This clone is available royalty-free through LNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Seq primer: -28M13 rev2 from Amersham
              High quality sequence stop: 419.
FEATURES
  source
    1..457
    /organism="Homo sapiens"
    /db_xref="GDB:380475"
    /db_xref="taxon:9606"
    /clone="IMAGE:490188"
    /clone_lib="Soares_pregnant_uterus_NbH9U"
    /sex="female"
    /dev_stage="adult"
    /lab_host="DH10B"
    /note="Organ: uterus; Vector: pT73-Pac; Site_1: Not I;
    Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
    oligo(dT) primer [5',
    AACTGGAAGATTCGCGCGCCCTTTTCTTTTCTTTTCTTTT 3'],
    double-stranded cDNA was ligated to Eco RI adaptors
    (Pharmacia), digested with Not I and cloned into the Not I
    and Eco RI sites of the modified pT73 vector. Library
    went through one round of normalization. Library
    constructed by M. Fatima Bonaldo."
BASE COUNT   119 a 82 c 112 g 142 t 2 others
ORIGIN
Query Match      18.9%; Score 385; DB 9; Length 457;
Best Local Similarity 99.8%; Pred. No. 1.9e-132;
Matches 435; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1287 GGAAGGATCCCGAGGCGAGATGATCAAGACCTGAATTCGACGCTTAAGACATGTC 1346
    |||||||
Db 1 GGAAGGATCCCGAGGCGAGATGATCAAGACCTGAATTCGACGCTTAAGACATGTC 60

```

QY 1347 CAGTGGCCCTTTGGAAAGCTGTGGGATCCAGTTACAGAGGAGGGGTATGTTCTCC 1405  
 |||||||  
 Db 61 CAGTGGCCCTTTGGAAAGCTGTGGGATCCAGTTACAGAGGAGGGGTATGTTCTCC 120  
 QY 1407 CAGCTGGGCAAGAGAGTGTATGCGCAATTCCTGCATAGCAGAAAGCTCCACCATT 1466  
 |||||||  
 Db 121 CAGCTGGGCAAGAGAGTGTATGCGCAATTCCTGCATAGCAGAAAGCTCCACCATT 180  
 QY 1467 TTTCTTTGATGTGTTTAAAGTCTACGTTCTCTATATAGAAACAGAGGTTCTGTCA 1526  
 |||||||  
 Db 181 TTTCTTTGATGTGTTTAAAGTCTACGTTCTCTATATAGAAACAGAGGTTCTGTCA 240  
 QY 1527 GCTCCTTGCTGCTGATGTGTGGAATGATGTTAGAGAAAGCATTTTCTTTC 1586  
 |||||||  
 Db 241 GCTCCTTGCTGCTGATGTGTGGAATGATGTTAGAGAAAGCATTTTCTTTC 300  
 QY 1587 TTTGAACCTTAAAGGTTCTATTTAAAGCAGACAGATTTTATATCATGAG 1646  
 |||||||  
 Db 301 TTTGAACCTTAAAGGTTCTATTTAAAGCAGACAGATTTTATATCATGAG 360  
 QY 1647 GATCTCTTTGTGTGAATACAGAGATGATGCTGATCCCTTAAAGAGTTTATGTCC 1706  
 |||||||  
 Db 361 GATCTCTTTGTGTGAATACAGAGATGATGCTGATCCCTTAAAGAGTTTATGTCC 420  
 QY 1707 CTGACTCTGGCTAAAA 1722  
 |||||||  
 Db 421 CTGACTCTGGCTAAAA 436  
 RESULT 14 543 bp mRNA linear EST 13-OCT-1998  
 A1193562  
 LOCUS ge0f07.x1 Soares fetal\_lung\_NbHL19W Homo sapiens cDNA clone  
 DEFINITION IMAGE:1744357 3', mRNA sequence.  
 A1193562  
 VERSION A1193562.1 GI:3744771  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 543)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 This clone is available royalty-free through LNL: contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seg primer: 40UP from G1BCO  
 High quality sequence stop: 472.  
 FEATURES  
 source  
 1. 543  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1744357"  
 /clone\_lib="Soares\_fetal\_lung\_NbHL19W"  
 /dev\_stage="19 weeks"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: Lung; Vector: p773D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - Oligo(dT) primer  
 (5'-TGTTACCAATCTGAGTGGAGCGCCGCAATTTTCTTTTCTTTT-3'),  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified p773 vector  
 (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M. Paloma Bonaldo. This library was constructed  
 from the same fetus as the fetal heart library, Soares  
 fetal heart NbHL19W."  
 BASE COUNT 181 a 114 c 84 g 164 t

ORIGIN  
 Query Match 18.2%; Score 372; DB 9; Length 543;  
 Best Local Similarity 99.6%; Pred. No. 1, le-127;  
 Matches 542; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1487 AGTCTACGTTCTCTATATAGAAACAGAGGTTCTGACGCTCTTGTGCTATGTC 1546  
 |||||||  
 Db 543 AGTCTACGTTCTCTATATAGAAACAGAGGTTCTGACGCTCTTGTGCTATGTC 484  
 QY 1547 GCTCGAAATGATGTTAGGAAAGCATTTTCTTCTTGAACCTTAAAGGTTCTTA 1606  
 |||||||  
 Db 483 GCTCGAAATGATGTTAGGAAAGCATTTTCTTCTTGAACCTTAAAGGTTCTTA 424  
 QY 1607 TTATTAAGCAGACAGATTTCCACATTTTATACATGAGAGATCTTCTTGTGTAATA 1666  
 |||||||  
 Db 423 TTATTAAGCAGACAGATTTCCACATTTTATACATGAGAGATCTTCTTGTGTAATA 364  
 QY 1667 CCAGAGTTGACCTGACCCCTTTAAAGAAAGTTTATGTCCTGACCTGCTAAATTTAT 1726  
 |||||||  
 Db 363 CCAGAGTTGACCTGACCCCTTTAAAGAAAGTTTATGTCCTGACCTGCTAAATTTAT 305  
 QY 1727 CTAAATTCAGATGCTTTTGTAGATGATGAGATTTTGTGACCCACATATTGGAGTTTC 1786  
 |||||||  
 Db 304 CTAAATTCAGATGCTTTTGTAGATGATGAGATTTTGTGACCCACATATTGGAGTTTC 245  
 QY 1787 TAGATTGAGTGAATGAGCAGAAAGGGCCATCTCCATTGAGATGATTAAGTGAACCAAC 1846  
 |||||||  
 Db 244 TAGATTGAGTGAATGAGCAGAAAGGGCCATCTCCATTGAGATGATTAAGTGAACCAAC 185  
 QY 1847 TAGTTCTCGAATCTTACAGAGAGAGGAATACACACAGAGAGCTGTGACATAGGAC 1906  
 |||||||  
 Db 184 TAGTTCTCGAATCTTACAGAGAGAGGAATACACACAGAGAGCTGTGACATAGGAC 125  
 QY 1907 TTGAAGACCAAGACTTTAAATTTGCGAGCTGCTCATGCTGAGATTATATACACGCG 1966  
 |||||||  
 Db 124 TTGAAGACCAAGACTTTAAATTTGCGAGCTGCTCATGCTGAGATTATATACACGCG 65  
 QY 1967 TCTTTCTATTGAGTTACAATCTATATTTTATGGAAGTTTAAATTAAGAAAAAATTTAC 2026  
 |||||||  
 Db 64 TCTTTCTATTGAGTTACAATCTATATTTTATGGAAGTTTAAATTAAGAAAAAATTTAC 5  
 QY 2027 AAGA 2030  
 |||||  
 Db 4 AAGA 1  
 RESULT 15 562 bp mRNA linear EST 27-MAR-2001  
 BG495857  
 LOCUS 602540248P1 NIH\_MGC\_59 Homo sapiens cDNA clone IMAGE:4671586 5',  
 DEFINITION mRNA sequence.  
 BG495857  
 ACCESSION BG495857.1 GI:13457373  
 VERSION BG495857  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 562)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNCM1486 row: 0 column: 11



FEATURES High quality sequence stop: 562.  
Location/Qualifiers  
source 1..562  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="4671586"  
/clone\_id="NIH\_MGC\_59"  
/tissue\_type="mucoepidermoid carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site: 1:  
Site1 (99ccgctcgcc); Site2: Site1 (99ccattatgccc);  
Double-stranded cDNA was prepared from cell line RNA. 5'  
and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dT(30)BN-3'  
sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dT(30)BN-3'  
(where B = A, C, or G and N = A, C, G, or T). Average  
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies  
contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC  
Library."  
BASE COUNT 167 a 122 c 142 g 131 t  
ORIGIN  
Query Match 18.2% Score 371; DB 12; Length 562.  
Best Local Similarity 99.4% Pred. No. 2.5e-127;  
Matches 521: Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 939 GACCTGTCCATTCTCCCTCTCTATGCTTAGAGTATCTGATCTCTGAAGTGGG 998  
DB 23 GACCTGTCCATTCTCCCTCTCTATGCTTAGAGTATCTGATCTCTGAAGTGGG 82  
OY 999 AGGACTCTGTCTTGAACCTGCTTGAATCTGCAAGTTTCTCCAGGCCACAAGC 1058  
DB 83 AAGACTCTGTCTTGAACCTGCTTGAATCTGCAAGTTTCTCCAGGCCACAAGC 142  
OY 1059 CTACAGCCACTCCATTAAGATGGCATTCAATGAAGTGAAGACAAGAACTTATCACC 1118  
DB 143 CTACAGCCACTCCATTAAGATGGCATTCAATGAAGTGAAGACAAGAACTTATCACC 202  
OY 1119 TGTGTACCTCTGTGATGAATCATCATTTGGGATCGGCAATGGCGCACATTAAT 1178  
DB 203 TGTGTACCTCTGTGATGAATCATCATTTGGGATCGGCAATGGCGCACATTAAT 262  
OY 1179 CCAATCCCACTTGAACCACTGAAGAAGAAAGATGGACTCAGATGCTGCACAA 1238  
DB 263 CCAATCCCACTTGAACCACTGAAGAAGAAAGATGGACTCAGATGCTGCACAA 322  
OY 1239 CCATGAAGAAGTGAAGTGTTCCTCCAGACTATTAACAAGAACCTAAAGGAGGATCCC 1298  
DB 323 CCATGAAGAAGTGAAGTGTTCCTCCAGACTATTAACAAGAACCTAAAGGAGGATCCC 382  
OY 1299 CAGGCGAGAATGATCAAGAGCTGAAGTCAAGCGTTTAAGAGACATGTCAGTGGCTTTG 1358  
DB 383 CAGGCGAGAATGATCAAGAGCTGAAGTCAAGCGTTTAAGAGACATGTCAGTGGCTTTG 442  
OY 1359 GAAAGTGTGGGATCCAGTTCAGAGGAGGAGGTATGTTGCTCCAGTCTGGGCAG 1418  
DB 443 GAAAGTGTGGGATCCAGTTCAGAGGAGGAGGTATGTTGCTCCAGTCTGGGCAG 502  
OY 1419 AGGAGTGTATCGGGAATTTCTGCATAGCAGAAAAGCTCCAC 1462  
DB 503 AGGAGTGTATCGGGAATTTCTGCATAGCAGAAAAGCTCCAC 546

Search completed: April 21, 2003, 22:17:05  
Job time : 3302 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 18:47:50 ; Search time 95.2977 Seconds  
(without alignments)  
8825.857 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 3575  
Sequence: 1 CTCGCCATAGATGCGCTCCG.....TTTACAGAGAAAAA 2041

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+np.model -DEV=x1h  
-Q/cgn2\_1/USPTO.spool/US09513151/runat\_15042003\_141144\_26380/app-query.fasta\_1.2446  
-DB=SPTRMBL\_21 -QFMT=fastan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosome2 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151.qcgn\_1.1\_125-@runat\_15042003\_141144\_26380 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-MAIN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: SPTRMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_prodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2256	63.1	435 4	Q96L45

2	2253	63.0	467	4	Q9H3H1	09h3h1 homo sapien
3	1550	43.4	326	4	Q9NXT7	09nxt7 homo sapien
4	1531	42.8	324	4	Q96FJ3	096fj3 homo sapien
5	1342	37.5	326	11	Q9D1H5	09d1h5 mus musculu
6	549	15.4	430	5	Q9G7G3	09g7g3 caenorhabdi
7	528.5	14.8	434	3	Q9U775	09u775 schistosach
8	493	13.8	466	10	Q9ZUX7	09zux7 arabidopsis
9	482.5	13.5	439	10	Q9S9Z6	09s9z6 oryza sativ
10	438.5	12.3	315	16	Q9R5S5	09r5s5 thermoaer
11	339	9.5	427	10	Q9S9W3	09s9w3 oryza sativ
12	338	9.5	318	10	Q9S6B0	09s6b0 arabidopsis
13	334.5	9.4	329	10	Q941D1	0941d1 arabidopsis
14	333.5	9.3	329	10	Q9LUG4	09lug4 arabidopsis
15	328.5	9.2	330	10	Q941D2	0941d2 arabidopsis
16	327.5	9.2	330	10	Q941C9	0941c9 arabidopsis
17	326	9.1	330	10	Q941C9	0941c9 arabidopsis
18	322.5	9.0	357	10	Q9CA35	09ca35 arabidopsis
19	321.5	9.0	357	10	Q941D3	0941d3 arabidopsis
20	313.5	8.8	330	10	Q91XB1	091xb1 arabidopsis
21	313.5	8.8	336	10	Q93WC9	093wc9 arabidopsis
22	302	8.4	463	10	Q9C5J6	09c5j6 arabidopsis
23	300	8.4	459	10	Q941C8	0941c8 arabidopsis
24	299	8.4	350	10	Q9S533	09s533 petunia hyp
25	283	7.9	303	16	Q9R5Z6	09r5z6 fusobacteri
26	276.5	7.7	379	5	Q8T1L1	08t1l1 dictyostell
27	267	7.5	342	10	Q9C6L1	09c6l1 arabidopsis
28	239.5	6.7	297	2	Q8VOS5	08vos5 methylobact
29	234	6.5	330	10	Q9RXY9	09rxy9 arabidopsis
30	135.5	3.8	706	10	Q9C6R8	09c6r8 arabidopsis
31	129	3.6	982	16	Q9U8F6	09u8f6 agrobacteri
32	128.5	3.6	820	16	Q9WY41	09wy41 thermotoga
33	125	3.5	447	17	Q9TUY9	09tuy9 methanopyru
34	117.5	3.3	559	16	Q99TX4	099tx4 staphylococ
35	117	3.3	383	5	Q9VCR9	09vcr9 drosophila
36	117	3.3	363	16	Q9SXB7	09sxb7 listeria in
37	116.5	3.3	952	10	Q9SXY7	09sxy7 arabidopsis
38	116.5	3.3	952	10	Q9M2Z6	09m2z6 arabidopsis
39	116	3.2	5171	4	Q8WKK9	08wkk9 homo sapien
40	116	3.2	26926	4	Q10466	010466 homo sapien
41	116	3.2	26926	4	Q8WZB3	08wzb3 homo sapien
42	116	3.2	34350	4	Q8WZ42	08wz42 homo sapien
43	115	3.2	325	16	Q8X151	08x151 clostridium
44	113	3.2	763	10	Q9LM99	09lm99 arabidopsis
45	112	3.1	747	2	Q53832	053832 staphylococ

## ALIGNMENTS

RESULT 1  
ID Q96L45 PRELIMINARY: PRT: 435 AA.  
AC Q96L45.  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE tRNA isopentenyl transferase (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2444833; PubMed=11560893;  
RA Lemaux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,  
RT "Regulation of Physiological Rates in Caenorhabditis elegans by a  
RT tRNA-Modifying Enzyme in the Mitochondria";  
RL Genetics 159:147-157(2001).  
DR EMBL: AY052768; AL14107.1; -  
DR InterPro: IPR002627; IPPT: 1.  
DR InterPro: IPR000822; Znf\_C2H2.  
DR Pfam: PF01715; IPPT: 1.  
DR ProDom: PD004674; IPPT: 1.

DR SMART; SM00355; ZnF\_C2H2; 1.  
 DR TIGR00174; miaa; 1.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 KW Transferase.  
 FT NON-TER 1  
 SQ SEQUENCE 435 AA; 48948 MW; 2279AE7C2D999FF1 CRC64;

## Alignment Scores:

Pred. No.:	2,15e-180	Length:	435
Score:	2256.00	Matches:	424
Percent Similarity:	99.77%	Conservative:	0
Best Local Similarity:	99.77%	Mismatches:	1
Query Match:	63.10%	Indels:	0
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q96L43 (1-435)

```

QY 29 GCACGAGAGATTCTGTGGGCACTGAGGGGCTGACAGGAGCCCTACCTTTGTA 88
   |||||
DB 1 AlaArgAlaValProValGlySerGlyLeuArgGlyLeuGlnArgThrLeuProLeuVal 20
QY 89 GTGATTCTGGGGCCAGGGGACCCGCAATCCAGCTGGCGTTGACGTAGGCCAGCG 148
   |||||
DB 21 ValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAlaLeuGlnLeuGlyGlnArg 40
QY 149 CTCGGCGGTGAGATGCTCAAGCCGTGACTCCATGACGCTGTATGAGGCTAGCATCATC 208
   |||||
DB 41 LeuGlyGlyLeuIleValSerAlaAspSerMetGlnValTyrGlnLysLeuAspIleIle 60
QY 209 ACCCAACAGGTTTCTGCCAGAGAGACAGAAATCTGCCGACACCATGATCAGCTTTGTC 268
   |||||
DB 61 ThrAsnLysValSerAlaGlnGlnGlnArgLysIleCysArgHisHisMetLysSerPheVal 80
QY 269 GATCCCTCTGTGACCAATTACACAGTGGTGACTCAGAAATAGACCAACTGCTGTGATT 328
   |||||
DB 81 AspProLeuValThrAsnTyrThrValValAspPheArgAsnArgAlaThrAlaLeuIle 100
QY 329 GAAGATATATTGGCCGAGACAAATCTCTATGTGTGGAGAGACCAATTATTCATT 388
   |||||
DB 101 GluAspIlePheAlaArgAspLysIleProIleValIleGlyLysThrAsnTyrTyrIle 120
QY 389 GAATCTCTCTCTGGAAGCTTCTTGCAATACCAGCCCCAGAGATGGGCACTGAGAAA 448
   |||||
DB 121 GluSerLeuLeuTyrLysValLeuValAsnThrLysProGlnLysMetGlyThrGlyLys 140
QY 449 GTGATGACCCGAAAGTGGAGCTTGAAGAGAGATGGTGTGTACTTACACAACCCCTTA 508
   |||||
DB 141 ValIleAspArgLysValIleLeuGlnLysGlnAspGlyLeuValLeuHisLysArgLeu 160
QY 509 AGCCAGGTGGACCCAGAAATGGCTGCCAGCTGCATCCATGACAAAGCAAGTGGCC 568
   |||||
DB 161 SerGlnValAspProGlnMetAlaAlaLysLeuHisProHisAspLysArgLysValAla 180
QY 569 AGGAGCTTGCAGATTTTGAAGAAACAGAAATCTCATAGTGAATTTCTCCATCTGCA 628
   |||||
DB 181 ArgSerLeuGlnValPheGlnGlnThrGlyLysSerHisSerGlnPheLeuHisArgGln 200
QY 629 CATACGGAAGAGTGGTGGTCCCTGAGAGTCTCTGAAGTCTCTTAACCTTGCATC 688
   |||||
DB 201 HisThrGlnGlnGlyGlyLysProLeuGlnGlyLysProLeuLysPheSerAsnProCysIle 220
QY 689 CTTTGGCTTCATGCTGACAGGACAGTTCAGATGAGCGCTGATAGAGGGGAGTGCAC 748
   |||||
DB 221 LeuTyrPheHisAlaAlaAspGlnAlaValAlaLeuAspGlnAlaValLeuAspArg 240
QY 749 ATGCTTGTGCTGGGCTCTTGGAGAGAACTAAGAGATTTTACAGAGCCTATATACAGAG 808
   |||||
DB 241 MetLeuAlaAlaGlyLeuLeuGlnGlnLeuLeuArgAspPheHisArgArgTyrAsnGlnLys 260
QY 809 AATGTTTCCGAAATAGCCAGAGACTATCAACATGCTATGTTTCCATCATCTTGCCTTCAAG 868
   |||||
DB 261 AsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLys 280

```

```

QY 869 GAATTTCACGAGTACCTGATCACTAGGAGAAATGCACACTGAGACTAGTACACGCTT 928
   |||||
DB 281 GluPheHisGlyLysThrLeuIleThrGlnGlyLysCysThrLeuGlnThrSerAsnGlnLeu 300
QY 929 CTAAGAAGAGACCTGCTCCATTTGTCCTCCCTGTCTATGCTTGAAGTATCTGATGTC 988
   |||||
DB 301 LeuLysLysGlyProGlyLysProIleValProProValTyrGlyLeuGlnValSerAspVal 320
QY 989 TCGAAGTGGAGAGAGTCTGTCTTGAAGCTGCTTGAATTCGTGCAAGTTTCAATCCAG 1048
   |||||
DB 321 SerLysTyrGlnGlnLysSerValLeuGlnProAlaLeuGlnIleValGlnSerPheIleGln 340
QY 1049 GGGCCACAGCCTACAGCCACTCCATTAAGATGTCCTACAAATGAGCTGAGAACAGAGATA 1108
   |||||
DB 341 GlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnAlaGlnLysLysArg 360
QY 1109 AGTTATCACCTGTGTGACCTGTGTGATTCGAATCATCATTTGGGATGCGCAATGGCAGCG 1168
   |||||
DB 361 SerTyrHisLeuGlnCysAspLeuGlnCysAspArgIleIleIleGlyAspArgGlnTyrPheAla 380
QY 1169 CACATTAATTCCAATCCCACTTGAACCAACTGAGAGAAAGAAAGAGATTGCACTCAGAT 1228
   |||||
DB 381 HisIleLysSerLysSerHisLeuAsnGlnLeuLysLysArgArgLysLeuAspSerAsp 400
QY 1229 GCTGTCAACACATAGAAAGTCAAGTGTTCCTCCGACTATTAACAAGACCTTAAGGG 1288
   |||||
DB 401 AlaValAsnThrIleGlnSerGlnSerValSerProAspTyrAsnLysGlnProLysGln 420
QY 1289 AAGGATCCCGCAGGCGAGATGATCAAGAGCTGAAATGCAAGCGTT 1333
   |||||
DB 421 LysGlySerProGlyLysAsnAspGlnGlnLeuLysCysSerVal 435

RESULT 2
ID Q9H3H1 PRELIMINARY; PRT; 467 AA.
AC Q9H3H1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE tRNA isopentenyl pyrophosphate transferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20564178; PubMed=11111046;
RA Golovko A., Hjalms G., Silbon F., Nicander B.;
RT "Cloning of a human tRNA isopentenyl transferase.";
RL Gene 258:85-93(2000).
DR EMBL; AF074918; AAC31324.1; -.
DR InterPro; IPR002627; IPTT.
DR InterPro; IPR000822; ZnF_C2H2.
DR Pfam; PF01715; IPTT. 1.
DR Prodom; PD004674; IPTT. 1.
DR TIGR00174; miaa; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 467 AA; 52725 MW; 634469919D7F56A5 CRC64;

Alignment Scores:
Pred. No.: 3,91e-180 Length: 467
Score: 2253.00 Matches: 439
Percent Similarity: 94.22% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 1
Query Match: 63.02% Indels: 26
DB: 4 Gaps: 1

US-09-513-151-3 (1-2041) x Q9H3H1 (1-467)
QY 11 ATGGCGTCCGTCGCGGCTGACAGAGACTTCTGTGGCAGTGGGCTGAGGGCTGCGAA 70
   |||||
DB 1 MetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArgGlyLeuGln 20

```

```

QY 71 CGACCCCTACTCTGTAGATGATCTCGGGGCGACGGCAGCCGGAATCCAGCTGGC 130
    |||||||
Db 21 ArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThyGlyLysSerThrLeuAla 40
QY 131 TTGCAGCTAGGCGCAGCGGCTCGGCGGTAGATCGTACGGCTGATCCATGCAAGCTTAT 190
    |||||||
Db 41 LeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValSerAlaAspSerMetGlnValTyr 60
QY 131 CAGGCGCTAGACATCATACACCAACAAGTTTGTGCCCAAGACAGAGAATCTCCGGCAG 250
    |||||||
Db 61 GlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleCysArgHis 80
QY 251 CACATGATCAGCTTGTGATCCCTCTGTGACCAATATACAGTGTGATCCAGAAAT 310
    |||||||
Db 81 HisMetIleSerPheValAspProLeuValThrAsnTyrThyValValAspPheArgAsn 100
QY 311 AGAGCACTGCTGTATGATGAAGATATATTTGCCCGAGACAAATTCCTATGTTGGGA 370
    |||||||
Db 101 ArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIleValValGly 120
QY 371 GGAACCAATATATACATGTAATCTCTGCTGGAAGTTCTTGTCAATACCAAGCCCGAG 430
    |||||||
Db 121 GlyThrAsnTyrTyrIleGlnSerLeuLeuTyrPlysValLeuValAsnThrLysProGln 140
QY 431 GAGATGGCGCAGTGAAGAAGTATGACCGAAAGTGGACCTGAAAGAGGAGGATGCTT 490
    |||||||
Db 141 GlnMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysLysArgGlyLeu 160
QY 491 GTACTTTCACAAAGCGCTAAGCCAGGTGAGCCAGCAATGGTCCCAAGCTGCATCCACAT 550
    |||||||
Db 161 ValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeuHisProHis 180
QY 551 GACAAACGCAAGTGGCCGAGAGCTTGCAAGTTTGAAGAACAAGATCTCTCATAGT 610
    |||||||
Db 181 AspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnThrGlyLysSerHisSer 200
QY 611 GAATTTCTCCATCGTCAACATACGAGAGAAGGTGGTGGCCCTGGAGCTGCTCGAAG 670
    |||||||
Db 201 GlnPheLeuHisArgGlnIleThrGlnGlnGlyGlyGlyProLeuGlyLysProLeuLys 220
QY 671 TTCTSTAACCTTGTGATCTTGTGGCTTCACTGTCGACAGGAGATTCTAGATGAGCCCTTG 730
    |||||||
Db 221 PheSerAsnProCysIleLeuTyrPheHisAlaAspGlnAlaValLeuAspGlnArgLeu 240
QY 731 GATTAAGAGGCTGATGACATCTGCTGCTGGCTTGTGAGAGACTAGAAGATTTTTCAC 790
    |||||||
Db 241 AspLysArgValAspAspMetLeuAlaIleGlyLeuLeuGlnGlnIleuArgAspPheHis 260
QY 791 AGACGCTATATCAGAGAATGTTCGGAATAATAGCCAGGATATCAACATGCTATCTTC 850
    |||||||
Db 261 ArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnIleGlyIlePhe 280
QY 851 CAATCAATTGGCTTCAAGGAATTTTCAAGTACCTGATCACTAGAGGAAATGCAACATG 910
    |||||||
Db 281 GlnSerIleGlyPheLeuGlnPheHisGlnTyrLeuIleThrGlnGlnLysCysThrLeu 300
QY 911 GAGACTAGTACCAGCTTTTAAAGAAAGG----- 940
    |||||||
Db 301 GlnThrSerAsnGlnLeuLeuLysGlyIleGlnAlaLeuLysGlnValThrLysArg 320
QY 941 -----CCTGGTCCCAT 952
    |||||||
Db 321 TyrAlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlnProIle 340
QY 953 GTCCGCGCTGTATGCTTGAAGATATCTGATGCTGCAAGTGGAGGAGCTGTATTTT 1012
    |||||||
Db 341 ValProGlnValTyrGlyLeuGlnValSerAspValSerLysTyrPrlGlnSerValLeu 360
QY 1013 GAACCTGCTTGAATGCTGCAAGTTCATGTCAGGCGCAAGGCTCAAGCAGCTGCA 1072
    |||||||
Db 361 GlnProAlaLeuGlnIleValGlnSerPheIleGlnGlnLysLysProThrAlaThrPro 380

```

```

QY 1073 ATTAAGATGCCATATACATGAACSTGAGAACAGAGAACTTATACCTGTGTGACCTGT 1132
    |||||||
Db 381 IleLysMetProTyrAsnGlnIleAsnGlnLysLysArgSerTyrHisLysCysAspLeuLys 400
QY 1133 GATCAATATCATATTGGGATGCGCAATGGCGAGCGCACATATAATCCAAATCCCACTTG 1192
    |||||||
Db 401 AspArgIleIleIleGlyAspArgGlnTyrPrlAlaAlaHisIleLysSerLysSerHisLeu 420
QY 1193 AACCAACTGAGAGAAAGAAAGAAATTTGACATGCTGCTCAACACCATACAGTCAAG 1252
    |||||||
Db 421 AsnGlnLeuLysLysArgArgArgLysAspSerAspAlaValAsnThrIleGlnSerGln 440
QY 1253 AGCTGTTTCCCGACACTATATACAAAGAACTTAAAGGAGAGATCCCGACGCGAGATGAT 1312
    |||||||
Db 441 SerValSerProAspHisAsnLysGlnPrlGlnLysGlySerProGlnLysAsnAsp 460
QY 1313 CAGAGCTGGAATGCAAGCTT 1333
    |||||||
Db 461 GlnGlnLeuLysCysSerVal 467

```

## RESULT 3

```

Q9NXT7 PRELIMINARY: PRT: 326 AA.
ID AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ20061 f1s, clone COL01383.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Oba-yashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isoqai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK000068; BAA90923.1; -.
DR InterPro: IPR002627; IPR1.
DR Pfam: PF01715; IPR1.
DR ProDom: PD004674; IPR1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 326 AA: 37435 MW: EAB3F0P9664B7ACE CRC64:

```

## Alignment Scores:

```

Pred. No.: 2,84e-121 Length: 326
Score: 1550.00 Matches: 298
Percent Similarity: 91.728 Conservative: 1
Best Local Similarity: 91.418 Mismatches: 1
Query Match: 43.368 Indels: 26
DB: Gaps: 1

```

US-09-513-151-3 (1-2041) x Q9NXT7 (1-326)

```

QY 434 ATGGGACATGAGAAATGATTTGACGCGAAAGTGGAGCTTGAAGAGAGATGCTGTGA 493
    |||||||
Db 1 MetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGlnAspIleLeuVal 20
QY 494 CTTCAACAGGCTAAGCCAGGCGAGCGAGCAAGAAATGGTCCAGAGCTGATCCATGAC 553
    |||||||
Db 21 LeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeuHisProHisAsp 40
QY 554 AAACGCAAGTGGCGAGAGCTTSCAAGTTTGAAGAACAAGATCTCTCATAGTGA 613
    |||||||
Db 41 LysArgLysValAlaArgSerLeuGlnValPheGlnGlnThrGlyLysSerHisSerGln 60
QY 614 TTTTCTCATGCTCAACATACGGAAGAGAGTGGTGGCTTGGAGAGCTCTGCAAGTTC 673
    |||||||
Db 61 PheLeuHisArgGlnHisThrGlnGlnGlyGlyGlyProLeuGlnGlyProLeuLysPhe 80

```

OY 674 TTAACCCCTGCATCCTTGGCTTACGTACGACGAGCAGTTCATGATGACGCTTGAT 733  
 DB 81 SerAsnProCysIleLeuTrpLeuHisAlaAspIleAlaValLeuAspGluArgLeuAsp 100  
 OY 734 AAGAGGCTGATGACATGCTTGGCTGCTGAGGCTTGGAGAACTAAAGATTTTTCACAGA 793  
 DB 101 LysArgValAspAspMetLeuAlaIleGlyLeuLeuGluIleuAlaGAspPheHisArg 120  
 OY 794 CGCTATATATGAGAAGATGTTTCCGAAATATGACGAGCTATGACATGCTTATCTTCAA 853  
 DB 121 ArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 140  
 OY 854 TCAATGCTTCAAGAAATTCACGAGTACCTGATCATGCTGAGGAAATACACACTGAG 913  
 DB 141 SerIleGlyPheLysGluPheHisGlyIleTrpLeuIleThrGluGlyLysCysThrLeuGlu 160  
 OY 914 ACTAGTAAACGCTTCTTAAAGAAAGCA-----CCTGCTCCCATTTGTC 940  
 DB 161 ThrSerAsnGlnLeuLeuLysGlyIleGluAlaLeuLysGlnValThrLysArgTyr 180  
 OY 941 -----CCTGCTCCCATTTGTC 955  
 DB 181 AlaArgLysGlnAsnArgTrpValLysAsnArgPheLeuSerArgProGlyProIleVal 200  
 OY 956 CCCCCTGTATGCTAGAGGTATCTGATGCTCGAAGTGGAGGAGCTGTCTTGA 1015  
 DB 201 ProProValTyrGlyLeuGluValSerAspValSerLysTrpGluGluSerValLeuGlu 220  
 OY 1016 CCTGCTCTTGAATCGTGAAGTTCATCCAGGCGCACAAAGCTTACAGCCACTCCATA 1075  
 DB 221 ProAlaLeuGluIleValGlnSerPheIleGlnGlyHisLysProThrAlaThrProIle 240  
 OY 1076 AAGATGCTATCATATGACGTGAGACAGACAGAAAGTATATCACTGTGACCTTGAT 1135  
 DB 241 LysMetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuCysAspLeuCysAsp 260  
 OY 1136 CGATCATCATGAGGATCCGAGTGGGAGCGCACATAAATCCAAATCCCACTTGAAC 1195  
 DB 261 ArgIleIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSerLysSerHisLeuAsn 280  
 OY 1196 CAACTGAAGAAAGAAAGATTTGACTGATGCTGTCAACACATAGAAAGTACAGAT 1255  
 DB 281 GlnLeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlnSerGlnSer 300  
 OY 1256 GTTCCCGACATATTAACAAGAACTAAAGGAGGATCCCGAGGCGCATGATCA 1315  
 DB 301 ValSerProAspHisAsnLysGluProLysGluLysGlySerProGlyGlnAsnAspGln 320  
 OY 1316 GAGCTGAATGACGCGTT 1333  
 DB 321 GlnLeuLysCysSerVal 326  
 RESULT 4  
 096FJ3 PRELIMINARY: PRT: 324 AA.  
 AC 096FJ3:  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-MAR-2001 (Tremblrel. 19, last sequence update)  
 DE Similar to tRNA isopentenyl pyrophosphate transferase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11) SEQUENCE FROM N.A.  
 RP TISSUE=UTERUS;  
 RC Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC010741; AAH10741.1;  
 DR InterPro: IPR002627; IPTT;  
 DR InterPro: IPR000822; znf\_C2H2.

DR Pfam: PF01715; IPTT: 1.  
 DR ProDom: PD004674; IPTT: 1.  
 DR SMART: SM00355; znf\_C2H2: 1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 DR Transferase.  
 SO SEQUENCE 324 AA; 37223 MW; 1E6835D7C09126A9 CRC64;  
 Alignment Scores:  
 Pred. No.: 1,11e-119 Length: 324  
 Score: 1531.00 Matches: 296  
 Percent Similarity: 91.10% Conservative: 1  
 Best Local Similarity: 90.80% Mismatches: 1  
 Query Match: 42.83% Indels: 28  
 DB: 4 Gaps: 2  
 US-09-513-151-3 (1-2041) x 096FJ3 (1-324)  
 OY 434 ATGGGACATGAGAAAGTATGATGACCGAAAGTGGAGCTTGAAGAGAGATGCTTGA 493  
 DB 1 MetClyThrGlyLysValIleAspArgLysValGluLeuGluLysGluAspGlyLeuVal 20  
 OY 494 CTTCACAACGCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGTGCATCCACATGAC 553  
 DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeuHisProHisAsp 40  
 OY 554 AANCGCAAAATGGCCAGAGCTTGCAGATTTTGAAGAAACAGAAATCTCATAGTGA 613  
 DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluIleThrGlyIleSerHisSerGlu 60  
 OY 614 TTTCTCCATGTCACATACGAGAAAGTGTGTGCTCCCTGGAGGCTCTGAAAGTTC 673  
 DB 61 PheLeuHisArgGlnHisThrGlnGluGlyGlyLysProLeuGlyLysProLeuLysPhe 80  
 OY 674 TTAACCCCTGCATCCTTGGCTTACGTACGACGAGCAGTTCATGATGACGCTTGAT 733  
 DB 81 SerAsnProCysIleLeuTrpLeuHisAlaAspGlnAla-----AspGluArgLeuAsp 98  
 OY 734 AAGAGGCTGATGACATGCTTGGCTGCTGAGGCTTGGAGGAGTTCATGATTCACAGA 793  
 DB 99 LysArgValAspAspMetLeuAlaIleGlyLeuLeuGluIleuAlaGAspPheHisArg 118  
 OY 794 CGCTATATGAGAAGATGTTTCCGAAATATGACGAGCTATGACATGCTTATCTTCAA 853  
 DB 119 ArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 138  
 OY 854 TCAATGCTTCAAGAAATTCACGAGTACCTGATCATGCTGAGGAAATACACACTGAG 913  
 DB 139 SerIleGlyPheLysGluPheHisGlyIleTrpLeuIleThrGluGlyLysCysThrLeuGlu 158  
 OY 914 ACTAGTAAACGCTTCTTAAAGAAAGCA-----CCTGCTCCCATTTGTC 940  
 DB 159 ThrSerAsnGlnLeuLeuLysGlyIleGluAlaLeuLysGlnValThrLysArgTyr 178  
 OY 941 -----CCTGCTCCCATTTGTC 955  
 DB 179 AlaArgLysGlnAsnArgTrpValLysAsnArgPheLeuSerArgProGlyProIleVal 198  
 OY 956 CCCCCTGTATGCTAGAGGTATCTGATGCTCGAAGTGGAGGAGCTGTCTTGA 1015  
 DB 199 ProProValTyrGlyLeuGluValSerAspValSerLysTrpGluGluSerValLeuGlu 218  
 OY 1016 CCTGCTCTTGAATCGTGAAGTTCATCCAGGCGCACAAAGCTTACAGCCACTCCATA 1075  
 DB 219 ProAlaLeuGluIleValGlnSerPheIleGlnGlyHisLysProThrAlaThrProIle 238  
 OY 1076 AAGATGCTATCATATGACGTGAGACAGACAGAAAGTATATCACTGTGACCTTGA 1135  
 DB 239 LysMetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuCysAspLeuCysAsp 258  
 OY 1136 CGAATCATATGAGGATCCGAGTGGGAGCGCACATGAGGCGCATGATCAATCCCACTTGA 1195  
 DB 259 ArgIleIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSerLysSerHisLeuAsn 278

```

Oy 1196 CAACGTGAAGAAAGAGATGCTGCTGCTCAACACCATGAAAGT 1255
    |||||||
Db 279 GlnLeuLysArgArgArgLeuAspSerAspAlaValAsnThrIleGluSerInser 298
Oy 1256 GTTCCCGACACTATACAAAGACCTTAAGGAGGATCCCGACGACGATGCA 1315
    |||||||
Db 299 ValSerProAspHisAsnLysGlnProLysGlnLysGlnSerProGlnAsnAspIn 318
Oy 1316 GAGCTGAATGCAGCGTT 1333
    |||||||
Db 319 GlnLeuLysCysSerVal 324

RESULT 5
Oy 09D1H5 PRELIMINARY: PRT: 326 AA.
AC 09D1H5:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE 2310075G14RIK protein (RIKEN CDNA 2310075G14 gene).
CN 2310075G14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RC MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aikawa K., Iwata M., Nishi K., Kiyosawa H., Kasukawa T., Salto R.,
RA Salto T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Salto R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschman W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bona M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK003556; BAB22853.1;
DR EMBL: BC019812; AAH19812.1;
DR MGI: MGI:1914216; 2310075G14RIK.
DR InterPro: IPR002627; IPT:
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPT: 1.
DR ProDom: PD004674; IPT: 1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 326 AA; 37191 MW; 454367A8B70DD1F0 CRC64;

```

## Alignment Scores:

```

Pred. No.: 7.63e-104 Length: 326
Score: 1342.00 Matches: 261
Percent Similarity: 83.74% Conservative: 12
Best Local Similarity: 80.06% Mismatches: 27
Query Match: 37.54% Indels: 26
DB: 11 Gaps: 1

```

US-09-513-151-3 (1-2041) x 09D1H5 (1-326)

```

Oy 434 ATGGCAGTGAAGAAAGTATTGATGACCGAAAGTGGAGCTTGAAGAGGATGCTTGTGA 493
    |||||||
Db 1 MetGlyThrGlyLysValValAlaAspArgLysValGluLeuGluLysGluAspLysGlu 20
Oy 494 GTTCAAAAGCGCTAAGCCAGGATGGAGCCAGAAATGGCTGCCAAGCTGCATCCAGATGAC 553
    |||||||
Db 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaValLeuHisProHisAsp 40
Oy 554 AAAGCAGGATGGCCAGAGAGCTTGCAAGTTTTGAAGAAGCAAGTCTCATAGTGA 613
    |||||||
Db 41 LysArgLysValAlaArgSerLeuGlnValProGluGluThrGlyIleSerHisSerGlu 60
Oy 614 TTTCATCGATCGTCAACATACAGAGAGAGTGGCTGCCCTGGAGGCTCTGTGAAGTTC 673
    |||||||
Db 61 PheLeuHisArgGlnHisAlaGluGluGluGlyProLeuGlyGlyProLeuArgPhe 80
Oy 674 TCTAACCTTGATGCTTGTGGCTTCATGCTGACAGGAGATTTAGATGAGCGCTTGAT 733
    |||||||
Db 81 ProAsnProCysIleLeuThrPheHisAlaAspGlnAlaValLeuAspGluArgLeuAsp 100
Oy 734 AAGAGGCTGATGATGATGCTGCTGCTGGCTCTGGAGAGACTTAAGATTTTCACAGA 793
    |||||||
Db 101 LysArgValAlaAspMetLeuAlaAlaGlyLeuLeuGluGluLeuArgGlyPheHisArg 120
Oy 794 CGCTATATCAGACAAATGTTTCGAAATATACCGAGACTATCAACATGCTATCTTCCA 853
    |||||||
Db 121 ArgTyrAsnLeuLysAsnIleSerGluAsnSerGluAspTyrGlnHisGlyIlePheGln 140
Oy 854 TCAATGGCTTCAAGAAATTTTCAGAGTACCTGATCACTGAGGAGAAATGCACTGGAG 913
    |||||||
Db 141 SerIleGlyPheLysGluPheHisGluTyrLeuThrThrGluGlyLysCysThrProGlu 160
Oy 914 ACTAGTACACCACTCTTAAGAAAGA----- 940
    |||||||
Db 161 ThrSerAsnGlnLeuLeuLysGlyIleGluAlaLeuLysGlnValThrLysArgTyr 180
Oy 941 -----CCTGGTCCCATGCTC 955
    |||||||
Db 181 AlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGluProSerVal 200
Oy 956 CCCCCTGTCTATGCTTACAGATGATGATGCTGCTGAGTGGAGAGAGTCTTGTGAA 1015
    |||||||
Db 201 ProProValTyrGlyLeuGluValSerAspValSerLysTyrGluGluSerValLeuGlu 220
Oy 1016 CCTGCTCTGAATGCTGCAAGTTTCATCCAGGCGGCAAGCCTTAAGCCTCCAA 1075
    |||||||
Db 221 ProAlaLeuAsnIleValGlnSerPheIleGlnGlnLysLysProThrAlaMetProVal 240
Oy 1076 AAGATGCCATATCAATGACCTGAGAAAGAGAGATTAACCTGTCGACCTGTGAT 1135
    |||||||
Db 241 LysMetAlaTyrAsnGluSerGluAsnLysLysArgSerTyrHisMetCysAspLeuGluAsp 260
Oy 1136 CGAATCATCATTTGGGATCGCAATGCGAGCGCAGCATTAATTCGAATCCCATGAA 1195
    |||||||
Db 261 ArgIleIleIleGlyAspArgGluTyrAlaAlaHisIleLysLysSerHisLeuHis 280
Oy 1196 CAACGTGAAGAAAGAGATTTGACCTGAGATGCTGCAACATGAAGATGAGATG 1255
    |||||||
Db 281 GlnLeuLysLysArgTyrArgLeuSerLeuAspAlaValSerAlaThrGlySerInser 300
Oy 1256 GTTCCCGACACTATACAAAGACCTTAAGGAGGATCCCGACGACGATGCA 1315
    |||||||
Db 301 AsnSerProAspCysAspProGluArgIleGluGluGluSerGlyGlnHisAsnGln 320

RESULT 6
Oy 09GYG3 PRELIMINARY: PRT: 430 AA.
AC 09GYG3:
Oy 09GYG3:

```



01-MAY-2000 (TReMBLrel. 13, last sequence update)  
 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
 tRNA isopentenyltransferase.  
 SPAC343.15.  
 Schizosaccharomyces pombe (fission yeast).  
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 Schizosaccharomycetales; Schizosaccharomycetaceae;  
 Schizosaccharomyces.  
 NCBI\_TaxID=4896;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL109739; CAB52278.1;  
 DR InterPro: IPR002627; IPT7.  
 DR Pfam: PF01715; IPT7.1.  
 DR ProDom: PD004674; IPT7.1.  
 DR TrEMBL: TIGR00174; miaA.1.  
 DR Transferase.  
 SO SEQUENCE 434 AA; 50121 MW; FCFD5101DF95F3AD CRC64;

Alignment Scores:  
 Pred. No.: 1,22e-35 Length: 434  
 Score: 528.50 Matches: 141  
 Percent Similarity: 50.00% Conservative: 79  
 Best Local Similarity: 32.05% Mismatches: 133  
 Query Match: 14.78% Indels: 87  
 Gaps: 19

US-09-513-151-3 (1-2041) x 09UT75 (1-434)

OY 80 CCTCTTGAGTGCATTCCTGGGCGCAGCGGACCGCAATTCACCGCTGCGTTCACCTA 139  
 DB 4 ProleucyValValIleGlyThrThrGlyAlaGlyLysSerAspLeuAlaValGlnLeu 23  
 OY 140 GCGCCGCGCTCGCGGTGAGATGCTCAGCGCTGACTGCATGAGTATGAGAGCCTA 195  
 DB 24 AlalysatgphneglyserGlnValIleasnAlaspsrmetGlnIleTyrArgGlyIle 43  
 OY 200 GACATCATCACCAACAGGTTTGTGCCACAGACAGAGATATGCCGCGCACCATGATC 259  
 DB 44 AsphTrIleThrAsnLysIleThrValGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 63  
 OY 260 AGCTTTGGATCCTCTGTGACCAATACAGTGTGAGTGTGACCTGACAAATAGACCACT 319  
 DB 64 SerPheLeuasn---PheAspLysGlnTyrSerValProGlnPheGlnArgAspAlaSer 82  
 OY 320 GCTCTGATTGAAGATATATTGTCGCCGAGACAAATCTTATTTGTTGGAGAGAACAT 379  
 DB 83 ArgValIleasprGlnIleHisSerGlnGlyLysIleProIleValIleGlyThrHis 102  
 OY 380 TATTACATGAATCTGCTGTGAAA-----GTTCTTGTG 415  
 DB 103 TyrTyrLeuGlnSerLeuLeuPheGlnAspThrThrLeuSerAlaIleAspLysLeuThr 122  
 OY 416 AAT-----ACCAAGCCCCAGACAGATGGCCACTGAGAAAGTATGACCGAATA 463  
 DB 123 AsnAspSerSerProSerLysProProHisProAspSerHis---IleLeuAsp----- 139  
 OY 464 GTGAGCTTGAAAAGAGATGCTGTGACTTACAAACCCCTAGACAGTGAACCA 523  
 DB 140 -----AspAspProSerAlaMetLeuSerTyrLeuLysLysIleAspPro 154  
 OY 524 GAAATGCTGCCAAGTGCATGCATGCATGACAAACGAAAGTGGCAGAGCTTGCACAT 583  
 DB 155 ValMetAlaGlnGlnThrPheIleProArgAspThrArgLysIleArgAspSerLeuGlnIle 174  
 OY 584 TTTCAGAACAGCAATCTCTCATAGTGAATTTCTCATCTGCTCAACATACGGAAGAGCT 643  
 DB 175 TyrPheHisThrGlyArgProProSerGlnIleTyrSerGlnGlnLysMetLysSerSer 194  
 OY 644 GGTGTCCTCCCTTGAGAGGCTGTGAGTTTCTTAACCTTGCATCCTTGTGCTCATCT 703

DB 195 -----GlySerLysLeuAsnArgTyrLysSer-----LeuIlePheThrAlaPheIle 209  
 OY 704 GACCAAGCAGTTTCTATGATGACGCTTGGATTAAGAGGTGATGATCATCTTGTCTGGC 763  
 DB 210 AspSerLeuValLeuMetProArgLysAspLysArgValAspLysMetLeuSerHisGly 229  
 OY 764 CTCTTGAGAGAACTAAGATTTTTCACAGACGCTATATCAACAATGTTCCGGAAT 823  
 DB 230 LeuValAspGlnIleLys-----SerMetLysSerLeuAlaGlnSer 243  
 OY 824 -----ACCCAGACTATCAACATGATCTTCAATCAATGCTTCAAGGAATTT 874  
 DB 244 GluLysPheSerProAspPheThrArgGlyIleThrGlnGlnIleGlyPheLysGlnIle 263  
 OY 875 CACGAGTACCTG-----ATCACTGAGGGAANAATGACACTGGAG----- 913  
 DB 264 MetProTyrPheGlnAlaProSerAspLysValPheAsnAspCys---LeuGlnArgMet 282  
 OY 914 -----ACTAGTAACCACTTCTAAGAAA----- 937  
 DB 283 LysValSerThrArgGlnTyrAlaLysSerGlnLysLysTyrPheIleGlnSerArgPheLeu 302  
 OY 938 -----GACCTGTGCTCCATTTGTCCTCCCTGTCTATGAC 970  
 DB 303 PrometCysLeuAlaGlnGlnAlaAspLeuSerProSerSerIleLeu-----PheSerThr 320  
 OY 971 TTAGAGTATCTGATGTCGCAAGTGGAGAGCTGTCTTGAACCTGCTTGAATC 1030  
 DB 321 ThrAsnThrThrAspLeuAsnAsnTyrPheGlnGlnVal---GluLysAlaCysArgVal 339  
 OY 1031 GTGCAAAATTTTCATCCAGGCGCACAAAGCTTACAGCCACTCCCAATTAACATCCATACAT 1090  
 DB 340 PheGlnTyrPhePheTyrAsnGlyAspAlaIleAla---ProSerAlaAspAspGlnHis 358  
 OY 1091 GAAGCTGAGAACAGAGAGATTAT----- 1114  
 DB 359 AlaPheGlnLysAlaAlaAspTyrLeuSerIleMetAsnGlyArgGlnSerGlnLysLys 378  
 OY 1115 ---CACCTGTGTGACCTGTGT---GATCGA-----ATCATCATTTGGGAT 1153  
 DB 379 LysPheValCysGlnGlnCysLeuAspLysArgGlyAspProPheThrValIleGlyGlu 398  
 OY 1154 CGCGAATGGCGACGACATTAATAATCCAAATCCCTGAACCAACGACAAAGAAAGA 1213  
 DB 399 AspAlaPheAsnValHisIleLysSerArgLysHisLysThrThrValArgArgLysLys 418

RESULT 8  
 ID 09ZUX7 PRELIMINARY: PRT: 466 AA.  
 AC 09ZUX7; Q8S059;  
 DT 01-MAY-1999 (TReMBLrel. 10, Created)  
 DT 01-JUN-2002 (TReMBLrel. 21, last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)  
 DE Putative tRNA isopentenylpyrophosphate transferase (TRNA  
 isopentenyltransferase).  
 GN ATG27760 OR ATPT72.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,  
 RA Shen M., Ronging C.M., Fraser C.M., Somerville C.R., Venter J.C.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.



RN [3] SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Town C.D., Kaul S.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Golovko A., Hajim G.;  
 RT "A tRNA isopentenyl transferase from Arabidopsis thaliana";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Takel K., Sakakibara H., Sugiyama T.;  
 RT "Identification of Genes Encoding Adenylylate Isopentenyltransferase, a  
 RT Cytokinin Biosynthesis Enzyme, in Arabidopsis thaliana.";  
 RL J. Biol. Chem. 0:0-0(2001).  
 DR EMBL: AC005824; AAC73024.2;  
 DR EMBL: AF109376; AAF00582.1;  
 DR EMBL: AB062609; BAB59042.1;  
 DR InterPro: IPR002627; IPT.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR Pfam: PF01715; IPT; 1.  
 DR ProDom: PD004674; IPT; 1.  
 DR SMART: SM00355; ZnF\_C2H2; 1.  
 DR TIGRfams: TIGR00174; miaa; 1.  
 KW Transferase.  
 SQ SEQUENCE 466 AA: 53098 MW: 59548DEB15FBDA CRC64;  
 Alignment Scores:  
 Pred. No.: 1,17e-32 Length: 466  
 Score: 493.00 Matches: 133  
 Percent Similarity: 50.10% Conservative: 108  
 Best Local Similarity: 27.65% Mismatches: 138  
 Query Match: 13.79% Indels: 102  
 DB: 10 Gaps: 18  
 US-09-513-151-3 (1-2041) x Q9ZUX7 (1-466)  
 QY 41 CTTGTGGCAGTGGCTCAGGGGCTCTGCAACGAGCCCTACCTCTTGTAGTAT 94  
 Db 6 6 Proserasnlglylglieglylglulysmetylslysalalysvalvalille 25  
 QY 95 CTCGGGGCCACGGCCGCAATCTCCAGCTTGGCAGCTAGCCGCGCTGGC 154  
 Db 26 MetclYprothrghlyserglylserlyslleualavalasleuaserHisphero 45  
 QY 155 GGTGAGATCGTACGCGCTGCTCCATGCTCAGTGTATGAGGCCCTAGACATCACCAC 214  
 Db 46 Valglutlleleasnalalaspalamecglniletyserglyleuaservalleuthrasn 65  
 QY 215 AAGTTTCTGCCCAAGACAGAGATCTGCCGCGCACCATGATCAGCTTTGTGATCT 274  
 Db 66 Lysvalthrvalasplglulnlyslvalprohishisileuleuglythrvaliser 85  
 QY 275 CTTTGACCAATTACAGAGTGGTGGACTTCAGAAATAGACCACTGCTGTGATGAAGAT 334  
 Db 86 Aspmet---gluphetrrhlaalargaspheargsphehrvalProleuilegluln 104  
 QY 335 ATATTCCCGCAGCAAAATTCCTATGTTGGTGGAGCAACATTTATCATGGAATCT 394  
 Db 105 lilevalseratrgasnhtstleprovalleuvalcylglythrhtstlyrtyllieglinala 124  
 QY 395 CTGCTCTGGAAGTTCTTGTCAATACCAAGCCGAGAGATGGGCACTGAGAAA----- 448  
 Db 125 Valvalserlyspheleuleuaspsalalaglulsp-----Thrglulglucyscs 142  
 QY 449 -----GTGATGACCGAAAGTGCAGCTGAA-----AAGGAG 481  
 Db 143 Alaaspvalalaservalalaspplaspmetvalvalgluservalpneglylrgasp 162  
 QY 482 GATGCTCTTGTACTTCAAAACGCGCTAAGCAGGTGAGCCAGAAATGGCTGCACACCTG 541

Db 163 AspleuserhislgltyrgrluleuleuLysgluleuasprovalalalalasnargyle 182  
 QY 542 CATCACATGACAACCAAGTGGCCAGAGCTTGCAGATTTCACAGAACGAAATC 601  
 Db 183 Hisproasnhtstlsergllysleasnlnlytleuserleuhtalaseratrglyval 202  
 QY 602 TCTCATAGCAATTTCTCCATGCTCAACATACGGAAGAGGTGGTCCCTGGAGGT 661  
 Db 203 Leuproserlys---leuhyrglnlylsthAlaglulsntrpdcylcystleasnAla 221  
 QY 662 CTTCTGAGTCTCTTAACCTTGCATCCTTGGCTTCATGCTGACAGGCACTTACAT 721  
 Db 222 Ser---Argphe---Aspryrcysleuileecyemecaspalaglulhrhralvalleuasp 239  
 QY 722 GACGCTTGGATAGAGGCTGATGACATGCTTGTCTGGCTGGCTTGGAGACTAACA 781  
 Db 240 Argtyvalglulnargvalalaspalamevalalaspalaglyleuueuaspgulvaltyr 259  
 QY 782 GATTTTCACAGACGCTATATACAGAAATGTTTCGAAATAGCCAGGACTTCACAAAT 841  
 Db 260 Asplletrylys-----Progllyalasptrythrarg 270  
 QY 842 GGTATCTTCCATCAATTTGGCTTCAAGAAATTCACAGATACCTG-----ATCACT 892  
 Db 271 GlyleuarglinsertleeglyvalArgglulphagluaspheuleuylstleuuser 290  
 QY 893 GAG-----GGAATGACACA---CTGGACACTATACCACTTCTTAAGAAAGCA 940  
 Db 291 Gluthrcysalagllyslleuthrserleuserasnaspysvalmetlysgluasn 310  
 QY 941 CTTGTCCTCCATT----- 952  
 Db 311 LeuathrglyslleuasnpheroLysaspaspysleuathrglulmetleugluAla 330  
 QY 952 ----- 952  
 Db 331 lileasprvalylsleuasnthrargargleuathrglulnlysatrgvalser 350  
 QY 953 ---GTCCCCCTGTCTATGCTTAGAGTATCTGAGTGC----- 968  
 Db 351 Argleuathrglulhrghlytrpasnlehtstlyrleaspladhrghlyrleu 370  
 QY 989 -----TCGAAGTGGAGAGCTGCTCTTGAACCTGCTTGAATCGTGAAC 1036  
 Db 371 SerlysserclulglusertrppasnAlaglvalvalylsptroalaserglullearg 390  
 QY 1037 AGTTTCATCCAG-----GCCACAGGCTACACCACTCCAAATTAAGATG 1081  
 Db 391 CysrphleuathrglulhrghlythrghlysergllyArgAsprthrserglylys 407  
 QY 1082 CCATACATGACGTGAGAAACAGAAAGTTATCACCTGTGTGACCTCTGT---GATCGA 1138  
 Db 408 -----SerlleuargaspheleuprthghlytrvalcylsAlaacysglyasnlys 425  
 QY 1139 ATCATGTTGGGTCGCGAATGGGCGGACCAATAAATCCAAATCCCACTTG----- 1192  
 Db 426 lleleuargllylthrghlyslutrrglunhtshisylsindllyrghtrhtshasgllyarg 445  
 QY 1193 -----AACCAACTGAGAAAGAAAGAAAGATTCAGTGCAGTGTCTGTC 1234  
 Db 446 ThrthrarghtsllysasnserglntthrtyllysasnarglulvalGlnlualaglulval 465  
 QY 1235 AAC 1237  
 Db 466 Asn 466  
 RESULT 9  
 ID 08S926 PRELIMINARY; PRT; 439 AA.  
 AC 08S926;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)





```

Db 229 -----AlaAspProIleAspAlaIleProArg 237
OY 542 CATCCACATGACAAACGCAAGTGGCCAGACCTTGCAGTTTTCAGAAACAGCAATC 601
Db 238 AspPro-----PheAlaAspAlaAspVal 245
OY 602 TCTCATAGTAAATTTCTCCATGCTCAACATACGAAAGAGTGGTCCCTCGCAGCT 661
Db 246 GilyTyrArgPro----- 249
OY 662 CCTCGAAGTCTCTAACCCTTCATCCCTTGGCTTCATGTCAGCAGCAGTTCAGAT 721
Db 250 AlaLeuArgPhePro---CysCysLeuLeuTrpValAspValAspAspValLeuAsp 268
OY 722 GAGCCCTGATAGAGGCTGATGATGATGCTTGTCTGCTGCTTGGAGAACTAGA 781
Db 269 GilyTyrLeuAspArgValAspAspMetValGlyGlyMetValGlyGlyLeuGly 288
OY 782 GATTTTCACAGACGCTATATCAAGAAATGTTTGGAAATAGCCAGACTATCAACAT 841
Db 289 GilyTyr-----PheAlaThrThrSerAlaSerGlyArgAlaSer-----HisAla 303
OY 842 GGTACTCTTCACATCAATGCTTGCCTTCAAGCAATTTCAAGCACTGATCAGCGAA 901
Db 304 GlyLeuGlyLysAlaIleGlyValProGlyLeuGlyAspTyrPhe-----AlaGlyArg 321
OY 902 TGCACACTGAGACTAGTATACAGCAGCTTCTAAG-----AAA 937
Db 322 LysSerLeuAspAlaIleAlaIleAspGlyIleLysAlaAsnThrArgValLeuAlaIleArg 341
OY 938 GGAACCTGCTCCATT-----GTCCCTCTGTCTATGCTTACAGTA----- 979
Db 342 GlnValGlyLysLeuArgArgMetAlaAspValTrpGlyTrpProIleArgArgLeuAsp 361
OY 980 -----TCTGATGCTCTGCAAG 994
Db 362 AlaThrAlaThrIleArgAlaArgLeuSerGlyAlaGlyArgAlaIleAlaIleAla 381
OY 995 TGGGAGAGCTGCTTCTTGAACCTGCTTGAATCGTGCAGAGTTTCATC 1045
Db 382 TrpGlyArgAspValArgGlyProGlyLeuAlaIleAlaMetArgGlnPheVal 398

RESULT 12
OYSB60 PRELIMINARY: PRT: 318 AA.
AC OYSB60:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Putative tRNA Isopentenyltransferase (Adenylate
DE Isopentenyltransferase) (Cytokinin synthase) (EC 2.5.1.27).
GN F22K18.150 OR ATG424650 OR ATPT4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OC eucosids II: Brassicales: Brassicaceae: Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Meyer M., Medler H., Medler E., Wambutt R., Hohnsels J., Mewes H.W.,
RA Bayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Medler H., Medler E., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

```

```

RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Takei K., Sakakibara H., Sugiyama T.;
RT "Identification of Genes Encoding Adenylate Isopentenyltransferase, a
RT Cytokinin Biosynthesis Enzyme, in Arabidopsis thaliana."
RL J. Biol. Chem. 0:0-0(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=WASSILEWSKIIJA;
RA Matsumoto M., Kakimoto T.;
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=WASSILEWSKIIJA;
RA Kakimoto T.;
RT "Arabidopsis cytokinin synthase."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL035356; CA22998.1; -
DR EMBL: AL161561; CAB79375.1; -
DR EMBL: AB062611; BAB59044.1; -
DR EMBL: AB061402; BAB59031.1; -
DR InterPro: IPR002627; IPTT.
DR Pfam: PF01715; IPTT; 1.
DR ProDom: PD004674; IPTT; 1.
KW Transiferase.
SQ SEQUENCE 318 AA: 36674 MW; BE8BD9F5E2E0F08 CRC64;

Alignment Scores:
Pred. No.: 1,02e-19 Length: 318
Score: 338.00 Matches: 97
Percent Similarity: 42.23% Conservative: 58
Best Local Similarity: 26.43% Mismatches: 90
Query Match: 9.45% Indels: 122
DB: 10 Gaps: 10

US-09-513-151-3 (1-2041) x OYSB60 (1-318)
OY 83 CTTGTAGTATCTCTCGGGGCGACGGCGCAATTCACGCTGCGCTTGCAGCTAGGC 142
Db 7 MetValValIleMetClyAlaThrGlySerGlyLysSerSerLeuSerValAspLeuAla 26
OY 143 CAGCGGCTCGCGGTAGATCGTCAGCGCTGATCTCATCAGTCTATAGAGCCTAGAC 202
Db 27 LeuHisPheLysAlaGlyIleIleAsnSerAspLysMetGlnPheTyrAspLysLeuLys 46
OY 203 ATCATCACCAACAGAGTTTCTGCGCAAGACAGAGATCTGCCGCGCACCATGATCAGC 262
Db 47 IleThrThrAsnGlnSerThrIleGlnAspArgArgGlyValProHisIleLeuGly 66
OY 263 TTGTGATCTCTTGTGACCAATTCACAGATGCTGACTTGAATAAGACCACTGCT 322
Db 67 GlnLeuAsnProGlnValIleAlaIleAlaGlnPheArgValMetLeuAlaGly 86
OY 323 CTGATGAAGATATATTTCGCCGACCAACAATTCATGTGTGGAGACCAATATAT 382
Db 87 AlaIleSerGlyIleThrGlnArgLysLeuProIleLeuAlaGlyLysSerAsnSer 106
OY 383 TACATGAATCTCTCTGGAAGATTTCTGCAATACCAACCCAGAGATGGCGACT 442
Db 107 TyrIleHisAlaLeuLeuAlaLys----- 114
OY 443 GAGAAAGTATGACCGAAAGAGTGAAGTGAAGAGAGATGCTTCTACTACAAA 502
Db 114 ----- 114
OY 503 CGCTTAGCCAGGTGACCCAGAAATGCTGCCAGACTCATCAGTACAGCAAGCAA 562
Db 115 -----SerTyrAspProGlnAsn-----TyrProPheSerAspHisLys 127
OY 563 GTGGCCAGAGAGCTTGCAGATTTTTCAGAAAGACAGAAATCTCATAGTGAATTTCTCAT 622

```



Db 246 GlnleuylslylleGlnargLeuHisLysLysTrpLysMetSerMetHisArgValAsp 265  
 Qy 977 GTATCTGATGTC-----TCGAAGTCGGAGAGACT 1006  
 Db 266 AlathrgluValPheLeuLysArgAsnValcIuGlucIlnasprgluaIatrgluAsnLeu 285  
 Qy 1007 GTTCTTGACCTGCTCTTGAATAGTGTCAAGTTTTCATCCAGGCCACCAAGCCTACAGCC 1066  
 Db 286 ValAlaArgProSerGluArgIleValAlaSplysPheTrpAsnAsn----- 301  
 Qy 1067 ACTCAATTAAGATGCCATACATGAATGAAGCTGAGACAAGAACTTATCAGCTGTGTGAC 1126  
 Db 302 -----AsnGlnleuLysAsnAspAspValcIuHis-Cys----- 312  
 Qy 1127 CTCGTGATCATCATCATTTGGGATCGCAATGGCAGCGGCACATTAAT 1178  
 Db 313 -----LeuAlaIasertyrgIyglYserGlyserArgAlaHisAsn 327  
 RESULT 14  
 Q9LUG4 PRELIMINARY: PRT: 329 AA.  
 AC Q9LUG4: 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE tRNA isopentenyl transferase-like protein (Adenylylate  
 DE isopentenyltransferase) (EC 2.5.1.27).  
 GN AtIPT7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID-3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA.  
 RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA.  
 RA MEDLINE-20277480; PubMed-10819329;  
 RX Nakamura Y.;  
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC  
 RT clones.";  
 RL DNA Res. 7:131-135(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA.  
 RA Takei K., Sakakibara H., Sugiyama T.;  
 RT "Identification of Genes Encoding Adenylylate Isopentenyltransferase, a  
 RT Cytochrome Biosynthesis Enzyme, in Arabidopsis thaliana.";  
 RL J. Biol. Chem. 0:0-0(2001).  
 DR EMBL: AB023036; BAB02782.1;  
 DR EMBL: AB062613; BAB59046.1;  
 DR InterPro: IPR002627; IPT.  
 DR Pfam: PF01715; IPT.1.  
 DR ProDom: PD004674; IPT.1.  
 DR Transferase.  
 SQ SEQUENCE 329 AA; 36979 MW; 66AFA61D3DE49AF8 CRC64;

## Alignment Scores:

Pred. NO.: 2,46e-19 Length: 329  
 Score: 333.50 Matches: 99  
 Percent Similarity: 38.94% Conservative: 56  
 Best Local Similarity: 24.87% Mismatches: 105  
 Query Match: 9.33% Indels: 138  
 DB: 10 Gaps: 9

US-09-513-151-3 (1-2041) x Q9LUG4 (1-329)

Qy 83 CTGTAGTATTCGGGGCCAGCGACCCGCAAAATCAGCTGGCTTCAGACTAGGC 142

Db 36 ValIlePheValMetGlyAlaThrGlySerGlyLysSerArgLeuAlaIleAspLeuAla 55  
 Qy 143 CAGCGCTCGCGGTGAGATGCTGACGCGCTACCTGCATGCGAGGCTTATGAAGCCCTAGAC 202  
 Db 56 ThrArgPheGlnGlyIleIleIleAsnSerAspLysIleGlnLeuTrpLysGlyLeuAsp 75  
 Qy 203 ATCAACCAACAAAGGTTTCCGCCAGACGAGCAAGTCTGCGCCGACCACTATGAC 262  
 Db 76 ValLeuThrAsnLysValThrProLysGlySarGlyValProHisLysLeuGly 95  
 Qy 263 TTTGTGATCTCTCTTGACCAATTAACACAGTGTGACTTCAGAAATAGACAATGCT 322  
 Db 96 ValPheAspSerGluAlaGlyAsnLeuThrAlaThrGlnTrpLysArgLeuAlaSerGln 115  
 Qy 323 CTGATTGACATATATTTGCCGAGACAAATTCCTATTGTTGGGAGACCAATTTAT 382  
 Db 116 AlaIleSerLysLeuSerAlaAsnAsnLysLeuProIleValAlaGlyLysSerAsnSer 135  
 Qy 383 TACATTGAATCTCTGCTCTGGAAGTCTTGTCAATACCAAGCCCGAGAGATGGCACT 442  
 Db 136 TyrIleGluAlaLeu----- 140  
 Qy 443 GAGAAAGTATGACCGAAAGTGAGCTGAAAGAGAGATGGTCTTACTTCACAAA 502  
 Db 140 ----- 140  
 Qy 503 CGCCTAAGCCAGGTGGACCCAGAAATGCTGCCAAGCTGCATCCACATGACAAACGCAA 562  
 Db 140 ----- 140  
 Qy 563 GTGGCAGAGACTTGCAAGTTTGAAGAACAGAAATCTCATAGTAATTTCTCAT 622  
 Db 141 -----ValAsnHisSerSer----- 145  
 Qy 623 CGTCAATACGAGGAAGAGTGTCCTCCCTGAGAGCTCTGAAAGTTCTTAACCT 682  
 Db 146 -----PheLeuLeuAsnAsnTrpAspCys 154  
 Qy 683 TGCATCCTTGGCTTACGCTGACCGACATTTCTATGAGCCCTGGATGAAGAGTGT 742  
 Db 155 CysPheIleTrpValAlaSerLysLeuProValLeuAsnSerPheValSerLysArgVal 174  
 Qy 743 GATACATGCTTGGCTGCGGCTCTGGAGAACATAAGATTTTCACAGCGCTTAAT 802  
 Db 175 AspArgMetMetGluAlaGlyLeuLeuGluValAlaGlyVal-----PheAsn 191  
 Qy 803 CAGAGAATGTTTGGAAATAGCCAGACTATCAACATGCTATCTTCAATCAATGGC 862  
 Db 192 ProLys-----AlaAsnTrpSerValcIyIleArgAlaIleGly 205  
 Qy 863 TTCAAGCAATTTACAGATACCTGATCAGTGAAGGAAATTCACACAGCACTAGTAAC 922  
 Db 206 ValProGluLeuHisGlyTrpLeuArgAsnGlnSerLeuValAspArgAlaThrLysSer 225  
 Qy 923 CAGCTTTTA-----AAGAAGAGACCTGTCCCATTTGTC----- 955  
 Db 226 LysMetLeuAspValAlaValLysAsnLysLysAsnThrGluIleLeuAlaCysArg 245  
 Qy 956 -----CCCTGTCTATGCTTACAG 976  
 Db 246 GlnleuLysLysIleGlnArgLeuHisLysLysTrpLysMetSerMetHisArgValAsp 265  
 Qy 977 GTATCTGATGTC-----TCGAAGTCGGAGAGACT 1006  
 Db 266 AlathrgluValPheLeuLysArgAsnValcIuGlucIlnasprgluaIatrgluAsnLeu 285  
 Qy 1007 GTTCTTGACCTGCTCTTGAATAGTGTCAAGTTTTCATCCAGGCCACCAAGCCTACAGCC 1066  
 Db 286 ValAlaArgProSerGluArgIleValAlaSplysPheTrpAsnAsn----- 301  
 Qy 1067 ACTCAATTAAGATGCCATACATGAATGAAGCTGAGACAAGAACTTATCAGCTGTGTGAC 1126

Db 302 -----AsnGlnLeuLysAsnAspAspValGlnHis-Cys----- 312

OY 1127 CTCGTGATCATCATTCATGGGATCGGATGGGACGCGACATTAAT 1178  
 Db 313 -----LeuAlaIaISerTyrGlyGlySerGlySerArgAlaHisAsn 327

RESULT 15

O94ID2 PRELIMINARY; PRT: 330 AA.

AC O94ID2: 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Cytokinin synthase.

GN ATP5.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MASSILEMSK1JA;

RA Matsumoto M., Kakimoto T.;

RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MASSILEMSK1JA;

RA Kakimoto T.;

RT Arabidopsis cytokinin synthase."

RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.

DR EMBL: AB061403; BAB59032.1; -

DR InterPro: IPR002627; IPTT.

DR Pfam: PF01715; IPTT; 1.

DR ProDom: PD004674; IPTT; 1.

SQ SEQUENCE 330 AA: 37408 MW: 0F440B0A4237CD02 CRC64:

Alignment Scores:

Pred. No.:	6,45e-19	Length:	330
Score:	328.50	Matches:	94
Percent Similarity:	36.368	Conservative:	53
Best Local Similarity:	24.548	Mismatches:	105
Query Match:	9.198	Indels:	131
DB:	10	Gaps:	7

US-09-513-151-3 (1-2041) x O94ID2 (1-330)

OY 20 GTGGCGGCTGCAGACGACTT-----CCTGTGGCGGCTGCAGGCGCTGCAGACGG 73  
 Db 5 MetThrAlaLeuArgIleValIleGlnProLeuSerLeuAsnPhenIleValAsnMetVal 24

OY 74 ACCCTACCT-----CTTGTAAGTATTCCTGGGGCCACGGGCGC 112  
 Db 25 AspValProPhePheArgArgLysAspLysValValPheValMetCylValThrGlyThr 44

OY 113 GCGAATCCAGCGCTGCAGACGAGCCAGCGCGCTGCAGGCGCTGCAGACCGT 172  
 Db 45 GlyLysSerArgLeuAlaIleAspLeuAlaThrArgPheProAlaIleValAsnSer 64

OY 173 GACTCATCAGAGTCTATGAGCGCTAGACATCATCACCACAGGTTCTGCCCAAGAG 232  
 Db 65 AspLysIleGlnValTyrIleGlyLeuAspIleValThrAsnLysValThrProGluGlu 84

OY 233 CAGAGAAATGTGGCGGACCAATGATCAGCTTGTGGATCTCTGTGACCAATTACAGA 292  
 Db 85 SerLeuGlyValProHisIleLeuLeuGlyThrValHisAspThrTyrGluAspPheThr 104

OY 293 GTGGGTGAGTTCAGAAATAGACAACTGCTGCATGATGAGATATATTTCCCGAGACAA 352  
 Db 105 AlaGluAspPheGlnArgIleAlaIleArgAlaValGluSerIleValGlnArgAspArg 124

OY 353 ATTCTATTGTGTGGGAGACCAATATTATCATTTGAATCTCTGCGAAAGTTCTT 412

Db 125 ValProIleIleAlaGlyLysSerAsnSerTyrIleGluAlaLeuVal----- 140

OY 413 GTCAATACCAAGCCCCAGAGAGATGGGACTAGACAAAGTATTCAGCGAAAGTGAGCTT 472

Db 141 -----AsnAspCysValAspPheArgLeuArgTyr 150

OY 473 GAAAGAGAGATGCTGTACTTACAAAGCGCTAAGCCAGGTGACCCAGAAATGGCT 532

Db 150 ----- 150

OY 533 GCCAAGCTGCATCCACATGACAAACGAAAGTGGCCAGAGCTTGCAAGTTTGAAGAA 592

Db 150 ----- 150

OY 593 ACAAGAAATCTCATAGTGAATTTCTCAATCGTCAACATACGGAAGAGTGGTGTCC 652

Db 150 ----- 150

OY 653 CTGGAGATCTCTGAAGTTCTCTAACCTTGCATCTTGGCTTCATGTCAGCAGCA 712

Db 151 -----AsnCysCysPheLeuThrValAspValSerArgPro 162

OY 713 GTTCTAGATGAGGCTTGGATAGAGAGGTGATGACATGCTTGTGCTGGCTTGGAG 772

Db 163 ValLeuHisSerPheValSerGluArgValAspLysMetValGlnMetCylLeuValAsp 182

OY 773 GAATAGAGATTTTACAGACGCTTAATACAGAAATGTTTGGAAATAGCCAGAC 832

Db 183 GluValArg-----ArgIlePheAspProSerSerAsp 194

OY 833 TTTCAACATGATCTTCCATCAATATGAGCTTCAAGAGATTTCAAGAGTCTGATCATC 892

Db 195 TyrSerAlaGlyIleArgArgAlaIleGlyValProGluLeuAspGluPheLeuArgSer 214

OY 893 GAGGAGAA---TGCACACTGGAGACTAGTAAACCAAGCTTGAAGAAAGACCTGTGCC 949

Db 215 GluMetArgAsnTyrProAlaGluThrThrGluArgLeuGlnThrAlaIleGluLys 234

OY 950 ATTGTCCCCCTGTCTATAGCTTACAGGATATGATGTCTCAAG----- 994

Db 235 IleLysGluAsnThrCysLeuLeuAlaCysArgGlnLeuGlnLysIleGlnArgLeuTyr 254

OY 994 ----- 994

Db 255 LysGlnThrLysThrAsnMetHisArgValAspAlaThrGluValPheLeuArgArgGly 274

OY 995 -----TGGAGAGAGTCTGTTTGAACCTGCTTGAATCGTGCAA 1036

Db 275 GluGluAlaAspGluAlaThrPheAsnSerValAlaHisProSerAlaLeuAlaValGlu 294

OY 1037 ACTTTCATC 1045

Db 295 LysPheLeu 297

Search completed: April 21, 2003, 18:55:40  
 Job time: 108.298 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:49:25 Search time 58.4237 Seconds

(Without alignments)  
5283.055 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575

Sequence: 1 CTGCCATTAAGATGCGCTCCG.....TTTACCAAGAAAAAAA 2041

Scoring table:

BLOSUM62	
Xgapop 10.0, Xgapext 0.5	
Ygapop 10.0, Ygapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus.n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09513151/runat\_15042003\_141145\_26460/app\_query.fasta.1.2446  
-DB=Published\_Applications\_AA -QFM=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09513151\_@cgn2\_1.1.15\_@runat\_15042003\_141145\_26460  
-NCPU=6 -ICPU=3 -NO\_XLPTX -NO\_MAP -LARGEOUTERT -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications AA:

1:	/cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
2:	/cgn2_6/ptodata/2/pubppaa/PTC_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubppaa/PCYUS_PUBCOMB.pep.*
8:	/cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
10:	/cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
12:	/cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
14:	/cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1074	30.0	222 10 US-09-764-853-586	Sequence 586, App
2	1074	30.0	222 10 US-09-764-853-784	Sequence 784, App
3	334.5	9.4	311 9 US-10-260-877-32	Sequence 32, Appl
4	301.5	8.4	301 9 US-09-738-626-5633	Sequence 5633, Ap

5	246	6.9	57 10 US-09-864-761-42725	Sequence 42725, A
6	167	4.7	218 10 US-09-881-752A-230	Sequence 230, App
7	116	3.2	26926 9 US-09-759-508B-2	Sequence 2, Appl1
8	115	3.2	628 10 US-09-828-447-12	Sequence 12, Appl
9	108	3.0	867 12 US-10-007-693-109	Sequence 109, App
10	106.5	3.0	620 10 US-09-815-242-11472	Sequence 11472, A
11	105.5	3.0	379 10 US-09-134-333-11	Sequence 11, Appl
12	103	2.8	671 9 US-09-792-630-41	Sequence 41, Appl
13	103	2.8	671 9 US-10-080-376-41	Sequence 41, Appl
14	103	2.8	671 9 US-09-953-351-41	Sequence 41, Appl
15	103	2.8	671 9 US-10-082-671-47	Sequence 47, Appl
16	103	2.8	671 9 US-10-097-100-41	Sequence 41, Appl
17	102.5	2.8	671 9 US-09-971-536-52	Sequence 52, Appl
18	101	2.8	671 9 US-09-792-630-43	Sequence 43, Appl
19	101	2.8	671 9 US-10-080-376-43	Sequence 43, Appl
20	101	2.8	671 9 US-09-953-351-43	Sequence 43, Appl
21	101	2.8	671 9 US-10-082-671-49	Sequence 49, Appl
22	101	2.8	671 9 US-10-097-100-43	Sequence 43, Appl
23	98.5	2.8	1044 9 US-10-027-591-2	Sequence 2, Appl1
24	98	2.7	1054 10 US-09-828-447-11	Sequence 11, Appl
25	96.5	2.7	430 9 US-10-117-015-6	Sequence 6, Appl1
26	96.5	2.7	577 9 US-09-738-626-5684	Sequence 5684, Ap
27	95.5	2.7	487 9 US-09-828-447-11	Sequence 66, Appl
28	95	2.7	506 9 US-10-157-223-7	Sequence 7, Appl
29	94	2.6	739 9 US-09-738-626-6988	Sequence 6988, Ap
30	94	2.6	854 10 US-09-815-242-5192	Sequence 5192, Ap
31	94	2.6	859 9 US-09-975-719-281	Sequence 281, App
32	93.5	2.6	752 10 US-09-815-242-13421	Sequence 13421, A
33	93.5	2.6	752 10 US-09-815-242-13661	Sequence 13661, A
34	93.5	2.6	1055 10 US-09-759-667A-3	Sequence 200, Appl
35	93	2.6	674 10 US-09-765-272-200	Sequence 2, Appl
36	93	2.6	2630 9 US-10-077-130-5	Sequence 5, Appl1
37	93	2.6	7968 9 US-10-077-130-5	Sequence 1444, Ap
38	92.5	2.6	531 10 US-09-925-300-1444	Sequence 4, Appl1
39	92.5	2.6	638 9 US-10-038-010-5	Sequence 8, Appl1
40	92.5	2.6	3075 10 US-09-938-275-5	Sequence 6, Appl1
41	91.5	2.6	684 10 US-09-732-180-8	Sequence 62, Appl
42	91.5	2.6	837 10 US-09-732-180-6	Sequence 5, Appl1
43	91.5	2.6	854 10 US-09-801-368-62	Sequence 62, Appl
44	91.5	2.6	854 10 US-09-732-180-5	Sequence 14, Appl
45	91	2.5	609 10 US-09-828-310-14	

#### ALIGNMENTS

RESULT 1  
US-09-764-853-586  
Sequence 586, Application US/09764853  
Patent NO. US20020090672A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OR INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P206  
CURRENT APPLICATION NUMBER: US/09/764,853  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 586  
LENGTH: 222  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (124)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (145)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-853-586  
Alignment Scores: 8.8e-94 Length: 222  
Pred. No.: 222



Score: 1074.00 Matches: 214  
Percent Similarity: 98.62% Conservative: 1  
Best Local Similarity: 98.17% Mismatches: 3  
Query Match: 30.04% Indels: 1  
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-586 (1-222)

QY 2 TGGCATAAGATGGCGTCGCGGGGCTGCACGAGCACTTCTGTGGCACTGGCTCAGG 61  
DB 1 CysHISLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20  
QY 62 GGCCTGCACAGGACCTTACCTTTGTAGTATCTCGGGGCGGCGGACCGGCAATCC 121  
DB 21 GilyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40  
QY 122 AGCGTGCGCTTGCAGTAGCGCCAGCGGCTCGCGGTAGATCGTCAGCGCTGCATCAG 181  
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60  
QY 182 CAGGCTATGAGAGCGCTTACATCATCACCACAGGTTTCTGCCAAGAGAGATC 241  
DB 61 GlnValTyrlGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80  
QY 242 TGGCGGACACATGATGACGCTTGTGATCCCTGTGACCAATACAGAGTGCGAGC 301  
DB 81 CysArgHISHisMetLieserPheValAspProLeuValThrAsnTyrlThrValAsp 100  
QY 302 TTCGAAATAGAGCAACTGCTGTGATTAAGATATATTCGCCGACACAAATTCCTATT 361  
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120  
QY 362 GTTGTGGGAGAACCAATTATTAATTGAATCTCTGCTGGAAGTTCTTGTCAATACC 421  
DB 121 ValValGly\*\*ThrAsnTyrlTyrlGlnSerLeuLeuTrpLysValLeuValAsnThr 140  
QY 422 AACCCCGAGAGATGGCGACTGAGAAAGATGATGACCGAAATGGGCGCTGAAAGAG 481  
DB 141 LysProGlnGlu\*\*GlyThrGlnLysValIleAspArgLysValGluLeuGlnLysGlu 160  
QY 482 GATGCTCTGTACTTTCACAAGCGCTAAGCAGGTGAGCCAGCAAAATGGCTGCCAGCTG 541  
DB 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180  
QY 542 CATCCACATGACAAAGCGCAAGTGGCCAGAGCTTGCAGTTTGAAGAAACAGAGATC 601  
DB 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGluThrGlyIle 200  
QY 602 TCTCATAGTGAATTTCTCATGCTCAACATAC-GGAAGAAGGTGGTGCCCT 654  
DB 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrpIleProSerPro 218

RESULT 2  
US-09-764-853-784  
Sequence 784, Application US/09764853  
Patent No. US20020090672A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P1206  
CURRENT APPLICATION NUMBER: US/09/764,853  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 784  
LENGTH: 222  
TYPE: PRN  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (124)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE  
LOCATION: (145)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-853-784

Alignment Scores:  
Pred. No.: 8-8e-94 Length: 222  
Score: 1074.00 Matches: 214  
Percent Similarity: 98.62% Conservative: 1  
Best Local Similarity: 98.17% Mismatches: 3  
Query Match: 30.04% Indels: 1  
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

QY 2 TGGCATAAGATGGCGTCGCGGGGCTGCACGAGCACTTCTGTGGCACTGGCTCAGG 61  
DB 1 CysHISLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20  
QY 62 GGCCTGCACAGGACCTTACCTTTGTAGTATCTCGGGGCGGCGGACCGGCAATCC 121  
DB 21 GilyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40  
QY 122 AGCGTGCGCTTGCAGTAGCGCCAGCGGCTCGCGGTAGATCGTCAGCGCTGCATCAG 181  
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60  
QY 182 CAGGCTATGAGAGCGCTTACATCATCACCACAGGTTTCTGCCAAGAGAGATC 241  
DB 61 GlnValTyrlGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80  
QY 242 TGGCGGACACATGATGACGCTTGTGATCCCTGTGACCAATACAGAGTGCGAGC 301  
DB 81 CysArgHISHisMetLieserPheValAspProLeuValThrAsnTyrlThrValAsp 100  
QY 302 TTCGAAATAGAGCAACTGCTGTGATTAAGATATATTCGCCGACACAAATTCCTATT 361  
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120  
QY 362 GTTGTGGGAGAACCAATTATTAATTGAATCTCTGCTGGAAGTTCTTGTCAATACC 421  
DB 121 ValValGly\*\*ThrAsnTyrlTyrlGlnSerLeuLeuTrpLysValLeuValAsnThr 140  
QY 422 AACCCCGAGAGATGGCGACTGAGAAAGATGATGACCGAAATGGGCGCTGAAAGAG 481  
DB 141 LysProGlnGlu\*\*GlyThrGlnLysValIleAspArgLysValGluLeuGlnLysGlu 160  
QY 482 GATGCTCTGTACTTTCACAAGCGCTAAGCAGGTGAGCCAGCAAAATGGCTGCCAGCTG 541  
DB 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180  
QY 542 CATCCACATGACAAAGCGCAAGTGGCCAGAGCTTGCAGTTTGAAGAAACAGAGATC 601  
DB 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGluThrGlyIle 200  
QY 602 TCTCATAGTGAATTTCTCATGCTCAACATAC-GGAAGAAGGTGGTGCCCT 654  
DB 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrpIleProSerPro 218

RESULT 3  
US-10-260-877-32  
Sequence 32, Application US/10260877  
Publication No. US20030021813A1  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Chovan, Linda E.  
APPLICANT: Hessler, Paul E.  
APPLICANT: Reich, Karl A.  
TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME  
TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION O  
FILE REFERENCE: 6565.US.PI  
CURRENT APPLICATION NUMBER: US/10/260,877

Oy	86	GTACGATGTTCTCCGGGCCACGGGCACCGGCAAAATTCSCAGCSTGGCGGTGCAGCTTAGGCCAG	145
Oy	87	.....	
Db	6	IlePheLeuMetGlyProThrAlaSerCylLysThrsAspLeuAlaIleGlnLeuAlaArgSer	25
Oy	146	C GGCTCCGGCGGTGAATCGTACAGCGCTGACTCCATTCGACAGCTATTGTGAAGCGCTGAGCATC	205
Db	26	GlnLeuProValGluValIleIleSerValAspSerAlaLeuIleTyrlTrpSgJmeMetAspIle	45
Oy	206	ATCACCAACAAGGTTTTTGCCCAAGACAGAGAAGATCTGCCGGCCACACATGATCAGCTTT	265
Db	46	GlyThrAlaLysProSerLysGluGluLeuValAlaLeuAlaProHisAlaArgLeuIleAspIle	65
Oy	266	GTGCACTCTCTTTGTGCACAAATTACACAGCTGGTGCATCCAAATAAGCAACACTGCTGTG	322
Db	66	LeuAspPro-----SerGluSerTrpSerAlaMetasnpPheAlaGaspAspAlaLeuAlaArgLu	84
Oy	326	ATTGAGATATATTATTTGCCCGGACAGACAATAATCTTATTGTGGGAGGAGAACCAATATTATAC	385
Db	85	MetaLasprIleThrAlaGlnGlyIleProLeuLeuValGluGlyTlrmethLeutyr	104
Oy	386	ATTGATCTCTGCTCTGGAAAAGTTCTTCAATTACCAAGCCCCACAGAGATGGGACACTGAG	444
Db	105	TyrlLysAlaLeu-----IleGluGlyLeuSerProLeuProSerAlaAspLu	120
Oy	446	AAGAGATTGACCCGAAAAAGTGCGCTTA-----AAGAGATGCGCTTGTA	492
Db	121	AsnIle-----ArgAlaGluLeuGluGlnLysAlaAlaGlnGlnGlyTrrAlaAla	133
Oy	494	CTTCAACAAACCGCTAAGCCAGGTGGACCAGCAAAATGCGTCCCAAGCTGCATCCACATGAC	552
Db	138	LeuHISthngIuLeuAlaLysIleAspProIleSerAlaAlaAlaGlyIleAsnProSerLasp	151
Oy	554	AAACSCAAAGTGGCCAGAGAGCTTGCAAGTTTTTGAAGAACAAGCAATCTCTCATAGTGA	613
Db	158	SerGlnAlaGlyIleAsnAlaGalaLeuGluValAlaPheTyrlIleThngIlysSerLeuThglu	177
Oy	614	TTTTTSCATCGTCAACATACAGAGAGAGAGAGTGGTGGCTCCCTGGAGGCTCTGTGAAGTTC	677
Db	178	Leu-----ThrlGlnLysLysGluAlaLeu-----ProtyrAspPhe	190
Oy	674	TCATAACCGCTGCATCGTTGGCTTGCATCTGACAGCGAGCTTATAGATGACCGCTGGAT	733
Db	191	ValGlnPheAlaIle-----AlaProGlnAspArgHisValAlaLeuHISgluAlaArgIleGlu	200
Oy	734	AAGAGGTGGATGACATGCTTGTGCTGGCTGGCTTTGAGAGAACTAAGAGATTTTCACAGA	792
Db	209	GlnAtsrPheHISLysMetIleGluLeuSjrgPheGlnAlaGluValGluLysLeuTyrala	222
Oy	794	CSCTATATACGAADAATGTTTTGCGAAATAGCCAGAGACTATCAACATGATATCTTCCA	855
Db	229	Arg---GlyAspLeuAsnIleLysIleuProser-----IleArg	240

386 ATTGAACTCTGCTC-----T

```

Oy      86  GAGGAGATTTCCTGGGGGGCCGGGCGACCGGCAAAATCCAGCGGGGTTGTCACCTTGGCCAG  145
Db      5  IIEAlaValGlgIProthraIaserIaIysSerIaIeugIIElaIeulAhaIis  24
Oy      146  CGGCTCGGGCGGTGAGATCGTAGCGCGCTGACTGCATGACGGTCTATGAAGCGCTTACATC  205
Db      25  LysLeuAspGlgIValIaValaIasnaIasIspSerMetGlnLeuTyIysGlyMetAspIle  44
Oy      206  ATCCACCAACACGGTTTCGGCCCAACAGGAGGAATCTCCGCAACACATCATGACGTTT  265
Db      45  GLYTHrAlaIaIysLeuThrIaIgluIuaIrgIuGlyIIElaIaIhisGlnLeuAspAl  64
Oy      266  GTGAGTCCTCTTGTGACCAATTACACAGTGTGGACTCTCAGAAATAGAGCAACTGCTCTG  325
Db      65  TrpAsp---ValThGluThrIaIaSerValAlaArgPheGlnSerAspAlaIaValaIaAsp  83
Oy      326  ATTGAAGATATATTTTCCCGACAGACAATAATCTCATATTTGTGGGAGCAACAAATTATTAAC  385
Db      84  ValGluAspIIEIeMetSerArgGlyIysThrProIIEuValGlyGlySerMetLeuTyI  103
Oy      386  ATTGAATCTGCTCTC-----TGGAAAGTTCTTGCAATTACCAAGCCCCAGAGATGGGC  435

```

Db 104 ValGlnSerLeuValAlaAspAspTrp-----GlnPheProPro 115  
 QY 440 ACTAGAGAACTGATTGACCCAAAAGTGACCTTGAAAAGAGAT---GGTCTT---GTA 493  
 Db 116 ThrAspSerAlaValAlaArgPheGlnAlaArgLeuAlaArgLeuValAla 135  
 QY 494 CTTCCAAAGCGCTAACGCCAGGTGAGCCGAATGGCTGCAAGCTGCATCCATGAC 553  
 Db 136 LeuHisAlaGluLeuThrGlnLeuAspProGlnAlaAlaValIleGluSerAsp 155  
 QY 554 AAACGCAAGTGGCCAGGACCTTGCAAGTTTGGAGAACAAGCAATCTCATAGTGA 613  
 Db 156 ProArgArgThrValAlaArgAlaLeuGluValIleGluLeuThrGly----- 170  
 QY 614 TTTCTCATGCTGACACTAGGAGAAAGAGGTGGTCCCTTGAGAGCTCTGAGATTC 673  
 Db 171 -----GlnProPheGlnAlaSerGlnProProLysAspAlaProArgTrp 186  
 QY 674 TCTAACCTTGATCTTGGCTTCATGCTGACAGGACAGTTCATGATGAGCGCTTGAT 733  
 Db 187 GlyThrArg---IleIleGlyLeuLysThrThrProGluTrpLeuAsnProArgIleGlu 205  
 QY 734 AAGAGGCTGATGACATGCTTGGCTGCTGCGCTTGGAGAGACTA 778  
 Db 206 GlnArgThrAlaArgMetPheGlnGlyPheValAlaGluVal 220

## RESULT 5

US-09-864-761-42725  
 ; Sequence 42725, Application US/09864761  
 ; Patent No. US20020048763A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Penn, Sharon G.  
 ; APPLICANT: Rank, David R.  
 ; APPLICANT: Hanzel, David K.  
 ; APPLICANT: Chen, Wensheng  
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
 ; FILE REFERENCE: Aeomica-X-1  
 ; CURRENT APPLICATION NUMBER: US/09/864,761  
 ; PRIOR FILING DATE: 2001-05-23  
 ; PRIOR APPLICATION NUMBER: US 60/180,312  
 ; PRIOR FILING DATE: 2000-02-04  
 ; PRIOR APPLICATION NUMBER: US 60/207,456  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: US 09/632,366  
 ; PRIOR FILING DATE: 2000-08-03  
 ; PRIOR APPLICATION NUMBER: GB 24263,6  
 ; PRIOR FILING DATE: 2000-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/236,359  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670  
 ; PRIOR FILING DATE: 2001-01-30  
 ; PRIOR APPLICATION NUMBER: US 60/234,687  
 ; PRIOR FILING DATE: 2000-09-21  
 ; PRIOR APPLICATION NUMBER: US 09/608,408

;; PRIOR FILING DATE: 2000-06-30  
 ;; PRIOR APPLICATION NUMBER: US 09/774,203  
 ;; PRIOR FILING DATE: 2001-01-29  
 ;; NUMBER OF SEQ ID NOS: 49117  
 ;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 ;; SEQ ID NO 42725  
 ;; LENGTH: 57  
 ;; TYPE: PRT  
 ;; ORGANISM: Homo sapiens  
 ;; FEATURE:  
 ;; OTHER INFORMATION: MAP TO AL033527.25  
 ;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98  
 ;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98  
 ;; OTHER INFORMATION: EXPRESSED IN BT4/4, SIGNAL = 1.1  
 ;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4  
 ;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
 ;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4  
 ;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2  
 ;; OTHER INFORMATION: EST HUMAN HIT: BE242161.1, EVALUE 3.00e-21  
 ;; OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALUE 2.00e-04  
 US-09-864-761-42725

## Alignment Scores:

Pred. No.:	4,76e-15	Length:	57
Score:	246.00	Matches:	48
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.88%	Indels:	0
DB:	10	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-864-761-42725 (1-57)

QY 185 GCTATGAGAGCGCTTAGACATCATCCACACAGCTTCTGCCAAGACAGACATCTGC 244  
 Db 1 ValTylGluLysLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleCys 20  
 QY 245 CGGACACACATGATCGAGCTTGGATCCCTTGACCAATACAGAGTGGGACTTC 304  
 Db 21 ArGHisHisMetLieserPheValAspProLeuValThrAsnTyrThrValAlaAspPhe 40  
 QY 305 AGAATAGAGCAACTGCTGTGATT 328  
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48

## RESULT 6

US-09-881-752A-230  
 ; Sequence 230, Application US/09881752A  
 ; Patent No. US20020115078A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kleantous, Harold  
 ; APPLICANT: Al-Garawi, Amal  
 ; APPLICANT: Miller, Charles  
 ; APPLICANT: Tomb, Jean-Francois  
 ; APPLICANT: Oomen, Raymond P.  
 ; TITLE OF INVENTION: Identification of Polynucleotides  
 ; TITLE OF INVENTION: Encoding No. US20020115078A1el Helicobacter Polypeptides in  
 ; FILE REFERENCE: 06132/041002  
 ; CURRENT APPLICATION NUMBER: US/09/881,752A  
 ; PRIOR FILING DATE: 2001-06-15  
 ; PRIOR APPLICATION NUMBER: US 08/833,457  
 ; PRIOR FILING DATE: 1997-04-01  
 ; NUMBER OF SEQ ID NOS: 370  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 230  
 ; LENGTH: 218  
 ; TYPE: PRT  
 ; ORGANISM: Helicobacter pylori  
 US-09-881-752A-230

## Alignment Scores:

Pred. No.:	3.13e-07	Length:	218
Score:	167.00	Matches:	55

Percent Similarity: 50.25%  
 Best Local Similarity: 27.09%  
 Query Match: 4.67%  
 DB: 10  
 Gaps: 10

US-09-513-151-3 (1-2041) x US-09-881-752A-230 (1-218)

OY 350 AAAATCTATTGTGGAGAGACCAATTATCATGATCTCTGTGGAAGT 409  
 Db 7 Gtulleleullelevalglysersephetylreulysertilleleulugly 25  
 OY 410 CTGTCAATACCAAGCCCGAGAGATGGACACTGAGAAAGTATGACGAAAGTGAG 469  
 Db 26 --LeuserArgmetProlyserGlyuglu-----ValVallys 39  
 OY 470 CTGAAAAGAGATGCTTCTTACTTCAACAAGC-----CTAAGCAGGTGACCCA 523  
 Db 40 lIeGuArGluIleAlaThrLeuSerAsnProTyrIlePheLeuLysertilleaspro 59  
 OY 524 GAAATGCTCCAGCTCATCCACATGACAAACGCAAGTGGCCAGCTTGCAGTT 583  
 Db 60 AsnMetAlaPheLysIleHisProAsnSprThrTyrArgThHisLysAlaLeuGluIle 79  
 OY 584 TTGAAGAACAAGAAATCTCTCATAGTGAATTTCTCATGTCATCAACATACGGAAGAGT 643  
 Db 80 PheTyrAlaThrCysThrProProserGluTyrPheLys----- 92  
 OY 644 GGTGTCCTCCCTGGAGGCTCTGAG---TTCTTAACCTTGTGATCTTTGG---CTT 697  
 Db 93 -----AlaAsnProLysLysProPheGluHisAlaIleSerLeuPheAlaLeu 108  
 OY 698 CATGTCAGCAGGAGCAGTCTAGATAGCGCTTGATAGAGGCTGATGACATGCTTGT 757  
 Db 109 SerIleGluLysSerAlaLeuHisAsnIleLysArgAlaGlyThLysSmMetLeuHis 128  
 OY 758 GCTGGCTCTTGGAGAACTAAGAGATTTCACAGACGCTATTAATCAGAAAGTCTTCG 817  
 Db 129 SerGlyLeuValIleGluIleLysAlaLeuTyrThrGlnTyr-----Pro 143  
 OY 818 GAAATACCAAGACTATACATGATGTTCTTCATCAATGCTTCAAGAAATTTTAC 877  
 Db 144 LysAspSerGlnPro-----PheLysAlaIleGlyValLysGlu----- 156  
 OY 878 GAGTACCTGATCATCTGAGGAAATGACACTGAGC-----ACTAGT 919  
 Db 157 SerValLeuPheLeuGluLysThrLeuThrLeuLysGluLeuGluAlaIleThrSer 176  
 OY 920 AACCAAGCTT 928  
 Db 177 AsnThrMet 179

RESULT 7

US-09-759-508B-2

Sequence 2, Application US/09759508B

Publication No. US20020182599A1

GENERAL INFORMATION:

APPLICANT: Fishman, Mark C.

TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease

FILE REFERENCE: 00786/381002

CURRENT APPLICATION NUMBER: US/09/759, 508B

PRIOR FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: US 60/175,787

NUMBER OF SEQ ID NOS: 11

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 26926

TYPE: PRT

ORGANISM: Homo sapiens

US-09-759-508B-2

Alignment Scores:

Pred. No.: 0.232

Length:

26926

Score: 116.00  
 Percent Similarity: 33.21%  
 Best Local Similarity: 20.49%  
 Query Match: 3.24%  
 DB: 9  
 Gaps: 25

US-09-513-151-3 (1-2041) x US-09-759-508B-2 (1-26926)

OY 13 GGGCGCGTGGGGGCTGC-----ACGAGCAGTTCCGTGGG 48  
 Db 7981 GlyLeuProGlyProCysLysAspIleLysAlaSerAspIleThrLysSerCysLys 8000  
 OY 49 CAGTGGGCTCAGGGGCTGCACAGGACCTTACTCTTGTAGTATTCGGGGCCAGCGG 108  
 Db 8001 LeuThrTrpGluProProGluPheAsp-----GlyGly 8011  
 OY 109 CACCGCAATTCACGCTGGCTGCTCAGCTACGCGGCGCTGGCGGTGAG----- 160  
 Db 8012 ThrProIleLeuHisTyrValLeuGluArgGluAlaGlyArgGly-ThrTyrIlePr 8031  
 OY 161 -ATGTCAGCGCTGACTCATGACGCTCATAGAGCCTACGATCATC----- 208  
 Db 8031 ovalMetSerGlyGluAsnLysLeuSerThrValLysAspLeuIleProAsnGlyG1 8051  
 OY 209 -----ACCAACAAGTTCTGCCCCAAGACAGACAATCTGCCG 246  
 Db 8051 uTyrPhePheArgValLysAlaValAsnLysAlaIleGlyGlyGluTyrIleGluLeuLys 8071  
 OY 247 GCACCAATGATCAGCTTGTGATCTCTGTGACCAATTAACAGAGTGGTGGACTTACG 306  
 Db 8071 sAsnProValIleAla---GlnAspProLysGlnProProAspProProValAspValG1 8090  
 OY 307 A---AATAGAGCACTGCTGTGATGAATATTTGCGGAGCAAAATTCCTATGCT 363  
 Db 8090 uValHisAsnProThrAla-----GluAlaMetThrIleThrTrpLysProProLeuTyr 8108  
 OY 364 TGTGGAGGAGAAC-----AATTATTACATGATATCTGCTC----- 400  
 Db 8108 rAspGlyGlySerLysIleMetGlyTyrIleIleLysIleAlaLysGlyGluGluArg 8128  
 OY 401 -TGAAA-----GTCTTGTCAATACC---AAGCCCCAGA 432  
 Db 8128 gTrpLysArgCysAsnGluHisLeuValProIleLeuThrTyrThrAlaLysGlyLeuG1 8148  
 OY 433 GATGGCACTGAGAAGAGATTTGACCGAAAGTGGAGCTTGAAGAAGAGATGCTTGT 492  
 Db 8148 uGluGlyLysGlu-----TyrGlnPheArgValArgAlaGluAsnAlaIleGlyIleSe 8166  
 OY 493 ACTTCACAAGCGCTTACG-----CAGTTGACCCAGCAAGATGCTGCCAGCT 540  
 Db 8166 rGluProSerArgAlaThrProProThrLysAlaValAspProIleAspAlaProLysVa 8186  
 OY 541 GCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAACAAGCAAT 600  
 Db 8186 lIleLeuArgThrSerLeuGluValLysArgLyspGluIleAlaLeuAspAlaSerI1 8206  
 OY 601 CTGTCAATGTAATTTCTTCATCGTCAACATACGAGAAGAGTGTGCTCCCTTGGAGG 660  
 Db 8206 eSerGlySerProTyr----- 8211  
 OY 661 TCCTCTGAAGTTCCTTAACCTTGCAATCTTGGCTTCATGCTGACGAGGATTCCTAGA 720  
 Db 8212 -----ProThrIleThrTrpLysAspGluAsnValIleValPr 8225  
 OY 721 TGAGCGCTTGATAGAGGTGATGACATGCTT-----GCTGCTGGCTCTTGGAGAGA 774  
 Db 8225 oGluGluIleLysLysArgAlaAlaProLeuValArgArgArgLysGluValGlnG1 8245  
 OY 775 ACTAAGAGATTTTCAACAGACCTTATATCAAGAGATGTTTCGAAATATACCGAGACTA 834  
 Db 8245 uGluGluProPheValLeuProLeuThrGlnArgLeuSerIleAspAsnSerLysLysG1 8265  
 OY 835 TCAA-----CATGTATCTTCCATCAAT 858

```

Db      8265 yGluSerGlnLeuArgValArgAspSerLeuArgProAspHisGlyLeuTyrMet----- 8283
OY      859 TGCGTTAAGCAATTTACAGACGATGACTGATCAGAGGAAATGCACTAGAGACTAG 918
Db      8284 ----lIeIyValGlnAsnAspHisGlylIeAlaIyAlaProCysThrValSerValle 8302
OY      919 TAACCACCTTCTAAGAAAGAGACTGGTCCCATTTGTCCTCCCTCTATATGCG-----TT 972
Db      8302 uAsPrThr-----ProGlyProProIleAsnPheValIleGlnAspIleAr 8317
OY      973 AGAGATATCTGATCTCTGCAAGTGGAGAGTGTCTTGAACCTGCTCTGAATGCT 1032
Db      8317 gLySThrSerValLeuCysIySTrPGLuProProLeuAspAspGlyGlySerGluIleI 8337
OY      1033 GCAAAATTCCATCCAGGCGCCAAAGCCTACAGCCATCTCCAAATAAGATGCCATCAATGA 1052
Db      8337 eAsnTyrThrLeuGlu---LysIyAspLysThrIyLsPro----- 8349
OY      1093 AGCGAGACAGACAGAAAGTTATACCTGTGTGACCTGTGTGATCATCATTTGGGGA 1152
Db      8350 ----- 8350
OY      1153 TCAGCAATGGCGACG-----CAC----- 1171
Db      8350 pSerGluTrpIleValIalThrSerThrLeuArgHisCysIySTySerValThrLysLe 8370
OY      1172 -ATAAATCCAAATCCACTTGAACCACTGACAGAAAGAAAGAAAGATTG----- 1219
Db      8370 uIleGluGlyLysGluTyrLeuPheArgValArgIalGlnAsnArpGheGlyProGlyPr 8390
OY      1220 -----GA 1221
Db      8390 oProCysValSerIySProLeuValAlaLysAspProPheGlyProProAspAlaProAs 8410
OY      1222 CTCAGATGCTGTCAACACCATAGAGATGAGTGTTCCTCCACACTATATAAGAAAC 1281
Db      8410 pLySTrProlValIalGlnAspValThrSerAsnSerMetLeuValLysTrpAsn---GluPr 8429
OY      1282 TAAAGGAGAGGATCCCA 1300
Db      8429 oLysAspAsnGlySerPro 8435

RESULT 8
US-09-828-447-12
: Sequence 12, Application US/09828447
: Patent No. US20020069432A1
: GENERAL INFORMATION:
: APPLICANT: COSTA E SILVA, OSMALDO DA
: APPLICANT: BOHNET, HANS J.
: APPLICANT: VAN THIELEN, NOCHA
: APPLICANT: CHEN, ROUYING
: APPLICANT: ISHITANI, MANABU
: TITLE OF INVENTION: SIGNAL TRANSDUCTION STRESS-RELATED PROTEINS AND METHODS
: FILE REFERENCE: 16313-0037
: CURRENT APPLICATION NUMBER: US/09/828,447
: CURRENT FILING DATE: 2001-08-20
: PRIOR APPLICATION NUMBER: 60/196,001
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 12
: LENGTH: 628
: TYPE: PRT
: ORGANISM: Physcomitrella patens
US-09-828-447-12

Alignment Scores:
Pred. No.: 0.0477 Length: 628
Score: 115.00 Matches: 101
Percent Similarity: 34.39% Conservative: 83
Best Local Similarity: 18.88% Mismatches: 169

```

```

Query Match: 3.22% Indels: 182
DB: 10 Gaps: 26
US-09-513-151-3 (1-2041) x US-09-828-447-12 (1-628)
OY      146 CGGCTCGCGCGTGATGATCGTACGCGGTAC-----TCCATGAG 184
Db      143 ArgArgGlyAlaArgValValIleGlnLeuAspLeuTrpProAspAspLysGlyGlyMetLys 162
OY      185 GTCTATGAAGCGCTAGACATCATCACACAGAGTTTGTGCCAA----- 229
Db      163 ValThrHisGly---AsnThrLeuThrAsnProValSerPheGlnLysCysValThrAla 181
OY      230 -----GACACAGAAATCTCCCGGACACCATGATCAGC 262
Db      182 lIeIyAsnAsnAlaPhePheThrSerGluTyrProValCys-----ValThr 197
OY      263 TTTGTGATCTCTTGGACCAATTACACAGTGTGACTCTCGAAATAGACCAACTGCT 322
Db      198 lIeGlnAspHisLeuThrSer-----GluLeuGlnGlyHisAlaIalGlu 212
OY      323 CTGATTGACAGATATATTGCGCCGACAGACAAATTCCTATTGTTGGAGAGAACCAATTAT 382
Db      213 lIeLeuGlnGlnIle-----LeuGlyAspAlaLeuTyr 223
OY      383 TAC-----ATGAATCTCTGCTCGGAA 406
Db      224 TyrProProThrThrAspAlaLeuValGluPheProSerProGluSerLeuLysArgLys 243
OY      407 GTTCTTGTCAATACCAAGCCCGAGAGATGGCAGTGAAGAAATGATGACCGAAAGTG 466
Db      244 lIeIleIleSerThrIySProProLysGluTyrLeuGlnLacysSerThrGlnLysLeu 263
OY      467 -----GACCTGAAAGAGAGATGCTTGTACTT--- 496
Db      264 AlawetGlnAsnArgAsnLeuValGluGlnIleGlnLysGlnAspLysLeuGlnGlnThr 283
OY      497 -----CACAAAGCGCTTAAGCCAGGTGAGACCGAAGAAATGGCTGCC 535
Db      284 ThrPheAlaProLeuGlnGlnAsnHisIleLeuGlyGlnAsnThrProSerLeuArgLys 303
OY      536 AAGCTGCATCCACATGACAAACGCAAGTGGCCGAGAGCTTGCAGCTTTTGAAGAAACA 595
Db      304 GluValIalValLeuSerGlnLysGlnMetSerThrProIalGlnLeuAsnSerArgSer 323
OY      596 GGAATCTCTATAGTGAATTTCTGCATCTGCAACATPACGAGAAAGTGCGTCCCTT 655
Db      324 ProSerAspLeuGlyGlnAlaThrSerThrArgTyrSerIySAsnAspGlyAsn--- 342
OY      656 GAGGCTCCTGAAAGTCTTAACCCCTGCATCTTGCGCTT-----CATGCT 703
Db      343 AspAsnProLysHisPheLysTyrAlaArgLeuIleThrIleArgLeuAlaLysHisAla 362
OY      704 GACCAGGCACTTCTAGATGACGCGCTTGAT-----AAGAGGCTG--- 742
Db      363 LysGlyThrSerMetGlnHisArgLeuGlnValAspGluSerValLysArgIleSerLeu 382
OY      743 GATGACATGCTTGTGCTGGCGCTTGGAGGAACTAAGACATTTTACAGACGCTTAAT 802
Db      383 SerCysLeuSerLeuGlnLysValIalGlnLysTrpProGlnAlaLeuValLysPheThr 402
OY      803 CACAGAGAT----- 811
Db      403 GlnLysAsnIleLeuArgValTyrProAlaAlaAsnArgValAsnSerSerAsnPheCys 422
OY      812 -----GTTTCGAAATATAGCCGACGATTCAA 838
Db      423 ProThrLeuAlaTrpAsnTyrGlyAlaGlnMetValAlaGlnAsnMetGlnGlyTyrGly 442
OY      839 CATGATATCTTCCAAATCAATGGC---TTCAGAGAAATTCACAGATCTGATCACTGAG 895
Db      443 LysGlnLeuSerProIalAlaPheGlyLysPheLys-----GlyAsn 455

```

```

RESULT 9
US-10-007-693-109
: Sequence 109, Application US/10007693
: Patent No. US20020146776A1
: GENERAL INFORMATION:
: APPLICANT: Bhacta, Ajay
: TITLE OF INVENTION: PROBST, Peter
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
: FILE REFERENCE: 210121.515C2
: CURRENT APPLICATION NUMBER: US/10/007, 693
: CURRENT FILING DATE: 2001-12-05
: NUMBER OF SEQ ID NOS: 157
: SEQ ID NO 109
: LENGTH: 867
: TYPE: PRT
: ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-109

```

Qy	17	TCCCTGGCGCCTGCACAGCAGTTCCTGTGGCAGTGGCCTCAGGGGCTGCACGCGACC	76
Db	583	SerIleArgAlaAlaArgVal-----GlyLeuSerAspProGlnArgPro	597

ORGANISM: Chlamydia trachomatis serovar D  
US-10-007-693-109

Pred. No.:	0.259	Length:	867
Score:	108.00	Matches:	72
Percent Similarity:	34.53%	Conservative:	34
Best local Similarity:	23.45%	Mismatches:	58
Query Match:	3.02%	Indels:	143
DB:	12	Gaps:	13

US-09-513-151-3 (1-2041) x US-10-007-693-109 (1-867)

Qy	17	TCCCTGGCGCCTGCACAGCAGTTCCTGTGGCAGTGGCCTCAGGGGCTGCACGCGACC	76
Db	583	SerIleArgAlaAlaArgVal-----GlyLeuSerAspProGlnArgPro	597

RESULT 10  
US-09-815-242-11472  
; Sequence 11472, Application US/09815242  
; Patent No. US2002006159A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlson, Karl L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard







```

US-09-513-151-3 (1-2041) x US-09-792-630-41 (1-671)
OY 1799 TCACCTCAATCTGAACTCCCAATATGTGGCTACAAATATCTTCACTGATCTACAAAGC 1740
DB 213 SerIleYsPheGlnThrMetValAsnTrpLeuGlySgluAsnArgValPheThrGluAsp 232
OY 1739 ATCGGAATATTAGATATTATTAGCCAG-----AGTCAGGACGATAAAGCTCT 1692
DB 233 LysTrpLysLeuValAspPheAsnGlnTrpThrLeuLeuSerSerHisSergLys 252
OY 1691 TTTAAGGATGACGATCAATCTG-----GTAATGACC 1659
DB 253 PheGln-----IleGlnSerAlaLeuLysLeuAlaIleTrpLysAlaThrAsnLeuValPro 271
OY 1658 ACAAGAAGATCTCATGATATAAATGTAATCTGTGCTTTTAATAATAGAACT 1599
DB 272 ThSerThrPheLeuLeuHisThrAspPheGlnGlnValMetCysIle----- 287
OY 1598 TTAAGGTTCAAGAAAAAATGCTTCTG-----AACTAC----- 1560
DB 288 -----LysAspAsnLysIleValLysLeuLeuLeuGlySgluAsnTrpAspProLeu 304
OY 1559 ATCATTTCCAGACATCAGCACACACAGAGAGTGACAGACGCTGTTCTATTATAG 1500
DB 305 LeuValAlGlnHisVal----- 310
OY 1499 AGAAGCTGAGACTTTAAACACATCAAAAGAAATGCTGAGACTTTCTGCTATGAC 1440
DB 311 -----LeuLysTrpIleAspLys----- 316
OY 1439 AGAATTCGCGATAGCACTCTTGGCCAGACTGGGAGACAAACTACC----- 1392
DB 317 -----LysCysGlyLysLysAsnThrLeuTrpPheTrp 327
OY 1391 ---CCTCCCTGCACTGATGATCCGACACACTTTCCAAAGGCCATGCACTGCTCT 1335
DB 328 GlyProProSer---ThrGly-----LysThrAsnLeuAlaMetAlaIle 341
OY 1334 AAAGCTGATTCGATCTTTGATCATCTGCGCTGGGATCCCTTCCCTTATGCTTC- 1277
DB 342 AlaLysSerValProValTrpGlyLysMetValAsnTrpAsnAsnGluAsnPheProPheAs 361
OY 1276 -----TTGTGATAGCTGGGGAACACACTGACTTTCATATGCTGTGA 1223
DB 361 nAspValAlaGlyLysSerLeuValValTrpAspGlnGlyIleLe----- 376
OY 1232 CAGCATCTGATCCCAATCTTCTCTTCTTCACTGATGTTCAAGTGGATTTGGATTTTA 1173
DB 377 -----LysSerThrIleValGlnAlaLysAlaIleLeuGlyGly 391
OY 1172 T-----GTGCGTGGCCATTCGCGATGCCATGCCAATGATGATTCGAT 1134
DB 391 nProThrArgValAspGlnLysMetArgLysSerValAlaValProGly----- 407
OY 1133 CACAGAGCTCACACAGCTGAATCTTCTT-----GTCTCAGCTCATTTATGCACTCT 1077
DB 408 -----ValProValAlaIleThrSerAsnGlyAspIleThrPheValValSergLys 425
OY 1076 TTATTGAGTGGCTGTAGCTTGTGGCTGGATGAACCTTTCAGACATTTCAAGACAG 1017
DB 425 nThrThrThrThrValHisAlaLysAlaLeuLysGluArgMet-----ValLysLeuAs 443
OY 1016 GTTCAACAACAGACTCTCCCA-----CTTCGAGACTTCAGATACCTCTAAGC 969
DB 443 nPheThrValArgLysSerProAspMetGlyLeuLeuThrGlnAlaAspValGlnGln 463
OY 968 CATGACAGAGGGGACAAATGGACAGAGCTCTTCTTAAAGAGTGGTACTACTGCCA 909
DB 463 rPleuThrTrpCysAsnAlaGlnSerTrpAspHisTrpGluAsnTrpAlaIleAsnTrp 483
OY 908 GTGTGCATTTTCCCTCAGTATCAGTACTGTAATTCCTTGAACCAATTGATTGGA 849
DB 908 GTGTGCATTTTCCCTCAGTATCAGTACTGTAATTCCTTGAACCAATTGATTGGA 849

```

```

DB 483 hrPheAspPheProGlyIle-----AsnAlaAspAlaLeuHisProAspLeuGln 500
OY 848 AGATACCATGTGATAGTCTGCTGATATTTCCGAACATTTCTGTATTATAGCTGTGT 789
DB 500 hrThrPro-----IleValThrAspThr----- 507
OY 788 GAAATCTTGTAGTCTCCCAAGAGAGCCACAGACATGATCATCCACCTTATGCCA 729
DB 508 -----SerIleSerSer-----GlyGlyGlnSerSergGlnLeuSerg 522
OY 728 AGCGCTCA-----TCAGAACCTGCTGTCAGCATGAAGCCAAAGATGC 684
DB 522 LysSerSerPhePheAsnLeuIleThrProGlyAlaTrpAsnThr----- 536
OY 683 AAGGTTAGAACTTCAGAGACCTCCAAAGGAGACCACTTCTCCGTATGTTGAC 624
DB 537 -----GluThrProArgSerSerThrPro-IleProGlyThrSer 549
OY 623 GATGAGAAATTCATATGAGATTCCTGTTCTTCAAAACTTGCAGACCTGCTG- 567
DB 550 SerGlyGlnSerPheValGlySerSerValSerSergIleValAlaIleAsnTrpGln 569
OY 566 -----CCACTTTC-----GTTGTCATGTGATGCA 540
DB 570 GluAlaPheTrpThrProLeuAlaAspGlnPheArgGlnLeuLeuValGlyValAspTrp 589
OY 539 GCTTGGCAGCATTTCTGGCTGACCTGCTTACGCTTGTGATGATGACAAAGCATCT 480
DB 590 ValTrpAsp-----GlyValAlaArgGlyLeuProValLysCysValGln----- 603
OY 479 CTTTTCACCTCCACTTTTGGTGCATCATCTTTCACAGGCCA-----TCTCCT 429
DB 604 -----HisIleAsnAsnSergLysGlyGlyLeuGlyLeuGlyProHisCysIleAsnVal 621
OY 428 GGGCTTGTGATGACAAAGACTTTCAGAGACAGATTCAAATGATATATGCG 375
DB 622 GlyAlaTrpTrp-----AsnGlyTrp 628

RESULT 13
US-10-080-376-41
; Sequence 41, Application US/10080376
; Patent No. US20020172968A1
; GENERAL INFORMATION:
; APPLICANT: DaiIyut. Basell I.
; TITLE OF INVENTION: BIOCHIPS COMPRISING NUCLEIC ACID/PROTEIN CONJUGATES
; FILE REFERENCE: A-70295-2/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/080,376
; CURRENT FILING DATE: 2000-02-19
; PRIOR APPLICATION NUMBER: US 09/792,630
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 671
; TYPE: PRF
; ORGANISM: B19 virus
; US-10-080-376-41

Alignment Scores:
Pred. No.: 0.688 Length: 671
Score: 103.00 Matches: 116
Percent Similarity: 33.89% Conservative: 67
Best Local Similarity: 21.48% Mismatches: 170
Query Match: 2.82% Indels: 187
DB: 9 Gaps: 30

US-09-513-151-3 (1-2041) x US-10-080-376-41 (1-671)
OY 1799 TCACCTCAATCTGAACTCCCAATATGTGGCTACAAATATCTTCACTGATCTACAAAGC 1740
DB 213 SerIleYsPheGlnThrMetValAsnTrpLeuGlySgluAsnArgValPheThrGluAsp 232

```

```

OY 1739 ATCTGGAATAGATAATTTTAGCCAG-----AGTCAGGACATAAACTTCT 1692
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 LysTrpLysLeuValAspPheAsnGlnTrpThrLeuLeuSerSerHisSerGlySer 252
OY 1691 TTTAAGGATGACATCACTCTG-----GTAATTCACC 1659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 253 PheGln-----IleGlnSerIleLeuLysLeuAlaIleTyrLysAlaThrAsnLeuValPro 271
OY 1658 ACAAGAAGATCCATATATATAAATGTGCAATCTGCTGCTTTATATATAACT 1599
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 ThrSerThrPheLeuLeuHisThrAspPheGlnGlnValMetCysIle----- 287
OY 1598 TTAAGCTCAAAAGAAAAAATGCTTCTCTG-----AACTAC----- 1560
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 288 -----LysAspAsnLysIleValLysLeuLeuLeuGlySerGlnAsnTyrAspProLeu 304
OY 1559 ATCATTTCCAGACACATAGCCACACAGAGAGCTGACAGACCTGCTTTCTATTATAG 1500
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 305 LeuValIleGlyIleHisVal----- 310
OY 1499 AGAAGCTGAGACTTTAAACCAATCAACAAAGAAATGCTGGAGCTTTCTGCTATGCAG 1440
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 311 -----LeuLysTrpLysAspLys----- 316
OY 1439 AGAATTCGCGATAGCACTCTTTGCCAGACTGGAGACAAACATACC----- 1392
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 -----LysCysGlyLysLysAsnThrLeuThrPheTyr 327
OY 1391 ---CCTCCCTCCCTGAATCGATCCCAACCTTTCCAAAGGCCATGACATGCTCTT 1335
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 GlyProPheSer-----ThrGly-----LysThrAsnLeuAlaIle 341
OY 1334 AAACGCTGCACTTACGCTTGATCATCTGCTCCGCGAGATCCCTTCTAGTTC-- 1277
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 342 AlaLysSerValProValTyrGly-MetValAsnTrpAsnAsnGlnAsnPheProPheAs 361
OY 1276 -----TTTGTATAGTCTGGGAAACACTGCACTTCTATGCTGTGA 1233
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AspValAlaGlyLysSerLeuValValTrpAspGluGlyIle----- 376
OY 1232 CAGCATCTGACATCTTCTTTCTTTCTTCTGATGCTTCAAGTGGATTTGGATTGTA 1173
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 377 -----LysSerThrIleValGlnAlaIleAlaLysAlaIleLeuGlyGly 391
OY 1172 T-----GTGCGCTGCCCATTCGCGAGATCCCATTCATGATTCAT 1134
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 391 nProThrArgValAspGlnLysMetArgLysValAlaValProGly----- 407
OY 1133 CACAGAGCTCACAGAGTGAATCTTCTT-----GTCTCAGCTTCATGATGACATCT 1077
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 408 -----ValProValValIleThrSerAsnGlyAspIleThrPheValSerGlyAs 425
OY 1076 TTATTGAGAGCTGTAGCTGTGGCCCTGGATGAACATTTCACAGATTCACAGACAG 1017
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 425 nThrThrThrThrValHisAlaLysAlaLeuLysGluArgMet-----ValLysLeuAs 443
OY 1016 GTTCAAGACAGACTCTCCCA-----CTTGAGACATCAATACCTTAAGC 969
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 443 nPheThrValArgCysSerProAspMetGlyLeuLeuThrGlu-AlaAspValGlnGln 463
OY 968 CATAGACAGGGGACCATGGAGACAGCTCTTCTTTAGAGAGCTGTACTAGCTCCCA 909
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 463 rPleuThrTrpCysAsnAlaGlnSerTrpAspHisTyrGlnAsnTrpAlaIleAsnTyr 483
OY 908 GTGTGCATTTTCCCTACAGTACAGGTACTGCTGAATCTCTGAGACCAATTGATTGGA 849
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 483 nPheAspPheProGlyLe-----AsnAlaAspAlaLeuHisProAspLeuGln 500
OY 848 AGATACCATGTGATAGCTGGCTATTTCGGAACATCTCTTGATTATAGCTCTGT 789
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 nThrPro-----IleValThrAspThr----- 507
OY 788 GAAATATCTTAGTCTCTCAAGAGCCACAGACAGACATGTCATCCACCTTATGCA 729

```

```

Db 508 -----SerIleSerSerSer-----GlyGlyGlySerSerGlnGluLeuSerG 522
OY 728 AGCGCTCA-----TCTAGAACTGCTGTGACATGAAGCAAGACATGC 684
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 IuSerSerPhePheAsnLeuIleThrProGlyAlaTrpAsnThr----- 536
OY 683 AAGGTTAGAGAACTTCAGAGAGACCTCCAGGGGACCAACCACTCTCCGATGTGAC 624
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 537 -----GluThrProArgSerSerThrPro-IleProGlyThrSer 549
OY 623 GATGAGAAATTCATATAGAGATCTGTTTCTTCAAAAATGCAAGCTCTG-- 567
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 SerGlyGluSerPheValGlySerSerValSerSerGluValAlaIleAsnTrpGlu 569
OY 566 -----CAGCTTGC-----GTTGTCATGTGATGCA 540
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 570 GluAlaPheTyrThrProLeuAlaAspGlnPheArgGluLeuValGlyValAspTyr 589
OY 539 GCTTGACAGCATTTGCTGGTCCACCTGCTTAGCGTTTGAGATCAAGACCATCCT 480
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 590 ValTrpAsp-----GlyValArgGlyLeuProValCysCysValGln----- 603
OY 479 CTTTTCAGCTCCACTTTTGGTCAATCACTTTCTCAGTGCACA-----TCTCCT 429
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 604 -----HisIleAsnAsnSerGlyGlyGlyLeuGlyLeuGlyCysProHisCysIleAsnVal 621
OY 428 GGGCTTGATTCAGCAAGAACTTTCACAGACAGATTCATGTAATATGG 375
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 GlyAlaTrpTyr-----AsnGlyTyr 628

RESULT 14
US-09-953-351-41
? Sequence 41, Application US/09953351
? Publication No. US2003003643A1
? GENERAL INFORMATION:
? APPLICANT: Li, Min
? APPLICANT: Melander, Christian
? APPLICANT: Liu, Hong-Xiang
? APPLICANT: Jin, Cheng He
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE CONSTRUCTION AND USE OF FU
? FILE REFERENCE: A-70814/RFT/RMS/RMK
? CURRENT APPLICATION NUMBER: US/09/953,351
? CURRENT FILING DATE: 2001-09-14
? PRIOR APPLICATION NUMBER: US 60/232,960
? PRIOR FILING DATE: 2000-09-14
? NUMBER OF SEQ ID NOS: 56
? SOFTWARE: PatentIn, version 3.1
? SEQ ID NO 41
? LENGTH: 671
? TYPE: PRT
? ORGANISM: B19 virus
US-09-953-351-41

Alignment Scores:
Pred. No.: 0.688 Length: 671
Score: 103.00 Matches: 116
Percent Similarity: 33.89% Conservative: 67
Best Local Similarity: 21.48% Mismatches: 170
Query Match: 2.82% Indels: 187
Gaps: 30

US-09-513-151-3 (1-2041) x US-09-953-351-41 (1-671)
OY 1799 TCACCAATCTAGAACTCCCAATATGTGCTCACAAATACTAGTCATGTACAAAGC 1740
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 213 SerIleLysPheGlnThrMetValAsnTrpLeuGlyAsnArgValPheThrGluAsp 232
OY 1739 ATCTGGAATAGATAATTTTAGCCAG-----AGTCAGGACATAAACTTCT 1692
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 LysTrpLysLeuValAspPheAsnGlnTrpThrLeuLeuSerSerHisSerGlySer 252
OY 1691 TTTAAGGATGACATCACTCTG-----GTAATTCACC 1659
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

QY 1598 TTAAAGTTCAAGAAAAAATGCTTTCTG-----AACTAC----- 1560
Db 288 -----LysAspAsnLysIleValIysLeuLeuLeuGlnAsnTyrAspProLeu 304
QY 1559 ATCATTTCCAGACACATGACCCACACAGAGAGCTGACAAAGACTGCTTTCTATTATAG 1500
Db 305 LeuValIglGlnHisVal----- 310
QY 1499 AGAACGTGAGACTTTAAACCACATCAAAAGAAATGGTGGAGCTTTCTGCTATGCAG 1440
Db 311 -----LeuLysTrrIleAspLys----- 316
QY 1439 AGAATTCGCATAGCACCTTTGGCCAGACTGAGAGACAACATACC----- 1392
Db 317 -----LysCysGlyLysLysAsnThrLeuTrrPheTyr 327
QY 1391 ---CCTCCCTCCTGAAGTGGATCCCAACCACTTTCCAAAGGCCATGACATGTCTCTT 1335
Db 328 GlyProPheSer---ThrGly-----LysThrAsnLeuAlaMetAlaIle 341
QY 1334 AAACCGCTGATTTCAAGCTTTGATCATCTGCGCTGGGATCCCTTCCCTTAGGTTG-- 1277
Db 342 AlaLysSerValProValTyrGly-MetValAsnTrrPasnAsnGluAsnPheProPheAs 361
QY 1276 -----TTGTATAGTCTGGGGAAACACTGCTGACTTCTATGCTTGA 1233
Db 361 nAspValAlaGlyLysSerLeuValValTrrPasnGluGlyIleIle----- 376
QY 1233 CAGCATCTGAGTCAACTCTTCTTTCTTCAGTTGCTTCAAGTGGATTTGGATTGTA 1173
Db 377 -----LysSerThrIleValIgluAlaLysAlaIleLeuGlyGly 391
QY 1172 T-----GTCCGCTGCCCATTCGCGATCCCAATGATGATTCAT 1134
Db 391 nProThrArgValAspGlnLysMetArgLysMetArgValAlaValProGly----- 407
QY 1133 CACAGAGTCAACAGAGTGAATCTTCTT---GTCTCAGCTTCATGTATGGCATCT 1077
Db 408 -----ValProValIleThrSerAsnGlyAspIleThrPheValSerGlyAs 425
QY 1076 TTATTGAGTGGCTGAGCTTGTGCGCTGATGAAACTTTGCACGATTTCAAGAGCAG 1017
Db 425 nThrThrThrThrValHisAlaLysAlaLeuLysGluArgMet-----ValLysLeuAs 443
QY 1016 GTTCAAGACACACTCCCTCCA-----CTCGAGACATCAGATACCTCTAACG 969
Db 443 nPheThrValArgCysSerProAspMetGlyLeuLeuThrGlu-AlaAspValGlnGln 463
QY 968 CATGACAGGGGGAACATGGACACAGCTCTTCTTAGAAGCTGGTACTAGTCTCA 909
Db 463 rPLeuThrTrrPcysAsnAlaGlnSerTrrPasnHisTyrGlnAsnTrrPalalIleAsnTyr 483
QY 908 GTGTGCAATTTCCCTCAGTATGAGTACTCGTGAATTCCTTGAAGCCAAATGATTTGA 849
Db 483 nPheAspPheProGlyIle-----AsnAlaAspAlaLeuHisProAspLeuGln 500
QY 848 ACATACCATGTTGATAGTCTGCTATTTCGGAACATCTTCTGATTATAGCGTCTGT 789
Db 500 nThrPro-----IleValThrAspThr----- 507
QY 788 GAAATCTGCTAGTCTCCAGAGCCCGACAGCAAGCATGTCATCAACCTCTATGCA 729
Db 508 -----SerIleSerSer-----GlyGlyGlnSerSerGluGluLeuSerG 522
QY 728 AGCGCTCA-----TCTAGAACTGCTGTGACATGAAAGCAAGGATGC 684
Db 522 LysSerPhePheAsnLeuIleThrProGlyAlaTrrPasnThr----- 536
QY 683 AAGGCTTAGAGACTTCAGAGACCTCCAGGGGACACCACTTCTTCCGTATGTTGAC 624
Db 537 -----GluThrProArgSerSerThrPro-IleProGlyThrSer 549

```

```

QY 623 GATGAGAAATTCACATAGAGATTCCTGTTTCTCAAAAACTTGCAGCTCCTG-- 567
Db 550 SerGlyLysSerPheValGlySerSerValSerSerGluValAlaIleAsnTrrGlu 569
QY 566 -----CCACTTGC-----GTTTGTCAATGTGATGCA 540
Db 570 GluAlaPheTyrThrProLeuAlaAspGlnPheArgGluLeuValGlyValAspTyr 589
QY 539 GCTTGGACAGCATTTCTGGTCCACCTGCTTAGCGCTTGTGAAGTACAGACCATTCCT 480
Db 590 ValTrrPasn-----GlyValArgGlyLeuProValCysCysValGln----- 603
QY 479 CCTTTCAAGCTCCACTTTTGGTCATCATCTTCTCAGTGCAC-----TCTCCT 429
Db 604 -----HisIleAsnAsnSerGlyGlyGlyLeuGlnCysProHisCysIleAsnVal 621
QY 428 GGGCTTGTGATTGACAGAACTTCCAGACAGAGATTCATATATATGCG 375
Db 622 GlyAlaTrrPyr-----AsnGlyTrr 628

```

Search completed: April 21, 2003, 19:01:37  
 Job time : 91.4237 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 19:18:40 ; Search time 37.5 Seconds

(without alignments)  
8230.823 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCATACATGCGTCGCG.....TTTACAGAAAAA 2041

Scoring table:

OLIGO	Xgapop 60.0	Xgapext 60.0
	Ygapop 60.0	Ygapext 60.0
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 288829 seqs, 75613885 residues

Word size: 1

Total number of hits satisfying chosen parameters: 541640

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:  
-MODEL=frame+olindp.model -DEV=xlh  
-O=/cgn2\_1/USPTO/spool/US05095131/Runat\_15042003\_141249\_27326/app.query.fasta\_1.2183  
-DB=Published\_Applications\_AA -OFMT=fastan -SUFFIX=olindp.rapb -MINMATCH=0.1  
-LOOCL=0 -LOOEXT=0 -UNITS=dots -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=45 -DICALIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US0509513151 -ECGN\_1\_1\_13 -Runat\_15042003\_141249\_27326  
-NCPU=6 -ICPU=3 -NO\_XLPTHY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published Applications\_AA.\*

1:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCIT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/PCITUS_PUBCOMB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	123	18.6	222	10	US-09-764-853-586
2	123	18.6	222	10	US-09-764-853-784
3	48	7.3	57	10	US-09-864-761-42725
4	8	1.2	115	10	US-09-867-550-1444

	5	8	1.2	139	10	US-09-867-550-482	Sequence 482, App
	6	8	1.2	378	9	US-09-970-989-17	Sequence 17, App1
	7	8	1.2	378	9	US-09-982-598-156	Sequence 156, App
	8	8	1.2	378	9	US-09-989-293A-156	Sequence 156, App
	9	8	1.2	378	9	US-09-988-725-156	Sequence 156, App
	10	8	1.2	378	9	US-09-990-444-156	Sequence 156, App
	11	8	1.2	378	9	US-09-989-730-156	Sequence 156, App
	12	8	1.2	378	9	US-09-990-436-156	Sequence 156, App
	13	8	1.2	378	9	US-09-991-181-156	Sequence 156, App
	14	8	1.2	378	9	US-09-993-667-156	Sequence 156, App
	15	8	1.2	378	9	US-09-988-724-156	Sequence 156, App
	16	8	1.2	378	9	US-10-028-072-420	Sequence 420, App
	17	8	1.2	378	9	US-09-997-653-156	Sequence 156, App
	18	8	1.2	378	9	US-09-993-667-156	Sequence 156, App
	19	8	1.2	378	9	US-10-121-904-420	Sequence 420, App
	20	8	1.2	378	9	US-10-123-904-420	Sequence 420, App
	21	8	1.2	378	9	US-10-140-470-420	Sequence 420, App
	22	8	1.2	378	9	US-09-990-438-156	Sequence 156, App
	23	8	1.2	378	9	US-09-990-562-156	Sequence 156, App
	24	8	1.2	378	9	US-09-997-428-156	Sequence 156, App
	25	8	1.2	378	9	US-09-997-666-156	Sequence 156, App
	26	8	1.2	378	9	US-10-175-746-420	Sequence 420, App
	27	8	1.2	378	9	US-10-176-918-420	Sequence 420, App
	28	8	1.2	378	9	US-10-176-921-420	Sequence 420, App
	29	8	1.2	378	9	US-09-990-711-156	Sequence 156, App
	30	8	1.2	378	9	US-10-137-865-420	Sequence 420, App
	31	8	1.2	378	9	US-10-140-474-420	Sequence 420, App
	32	8	1.2	378	9	US-10-142-431-420	Sequence 420, App
	33	8	1.2	378	9	US-10-143-114-420	Sequence 420, App
	34	8	1.2	378	9	US-09-989-726-156	Sequence 156, App
	35	8	1.2	378	9	US-10-140-002-420	Sequence 420, App
	36	8	1.2	378	9	US-09-990-437-156	Sequence 156, App
	37	8	1.2	378	9	US-09-998-156-156	Sequence 156, App
	38	8	1.2	378	9	US-10-142-419-420	Sequence 420, App
	39	8	1.2	378	9	US-09-991-157-156	Sequence 156, App
	40	8	1.2	378	9	US-09-991-172-156	Sequence 156, App
	41	8	1.2	378	9	US-09-997-573-156	Sequence 156, App
	42	8	1.2	378	9	US-09-997-573-156	Sequence 156, App
	43	8	1.2	378	9	US-10-123-282-420	Sequence 420, App
	44	8	1.2	378	9	US-10-142-423-420	Sequence 420, App
	45	8	1.2	378	9	US-09-990-443-156	Sequence 156, App

#### ALIGNMENTS

RESULT 1  
US-09-764-853-586  
Sequence 586, Application US/09764853  
Patent No. US20020090672A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PJ206  
CURRENT APPLICATION NUMBER: US/09/764,853  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 586  
LENGTH: 222  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (124)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (145)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-853-586  
Alignment Scores:  
Pred. No.: 2.12e-112 Length: 222



```

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42725
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033527.25
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EST_HUMAN HIT: BE242161.1, EVALUATE 3.00e-21
; OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALUATE 2.00e-04
US-09-864-761-42725

Alignment Scores:
Pred. No.: 2,51e-38 Length: 57
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.26% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-864-761-42725 (1-57)
OY 185 GTCATGAGCCCTAGACATATCCACACAGGTTTGGCCACAGACAGAACTCTGC 244
      |||||||
Db 1 ValTyrgIugIleuAspIleIleThraSnlYValSerAlaGInGInGInArG1IleCys 20
OY 245 CGGACACGATGATCAGCTTGTGAGACCTCTTGACCACTTACAGCTGGAGCTTC 304
      |||||||
Db 21 ArgIshIshEcIleSerPheValAspProIeuValThrAsnTrYThValValAspPhe 40
OY 305 AGAAATAGAGCACTGCTCTGATT 328
      |||||||
Db 41 ArgAsnArGAlaThrAlaLeuIle 48

RESULT 4
; Sequence 1444, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1444
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: Wherein Xaa may be any one of Arg or Gly or Trp
US-09-867-550-1444
```

```

Alignment Scores:
Pred. No.: 70.2 Length: 115
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-867-550-1444 (1-115)
OY 811 ATCTCTGATTATAGCCTCTGTC 788
      |||||||
Db 2 IleuIeuIleIleAlaSerVal 9

RESULT 5
US-09-867-550-482
; Sequence 482, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cell
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 482
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-482

Alignment Scores:
Pred. No.: 69 Length: 139
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-867-550-482 (1-139)
OY 113 GGCAATTCACGCTGGCTTCAG 136
      |||||||
Db 14 GlYlYserThrIleuAlaLeuGln 21

RESULT 6
US-09-970-989-17
; Sequence 17, Application US/09970989
; Patent No. US20020156262A1
; GENERAL INFORMATION:
; APPLICANT: LEUNG, David W.
; APPLICANT: ADOUREL, Daniel
; APPLICANT: HOLLENBACK, David
; TITLE OF INVENTION: MAMMALIAN LYOPHOSPHATIDIC ACID ACYLTRANSFERASE
; FILE REFERENCE: 077319/0151
; CURRENT APPLICATION NUMBER: US/09/970,989
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/215,252
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: US 08/618,651
; PRIOR FILING DATE: 1996-03-19
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
```

```
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-989-17

Alignment Scores:
Pred. No.: 62.8      Length: 378
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 1.21%          Indels: 0
DB: 9                Gaps: 0

US-09-513-151-3 (1-2041) x US-09-970-989-17 (1-378)

QY 669 AGTTCTCTAACCCTGCATCTT 692
    |||||
Db 334 SerSerLeuThrLeuAlaSerPhe 341

RESULT 7
US-09-992-598-156
; Sequence 156, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerbtsen, Hanspeter
; APPLICANT: Goddard, Audrey E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730PIC20
; CURRENT APPLICATION NUMBER: US/09/992,598
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
```

```
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
```



;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07

;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09  
  
Alignment Scores:  
Pred. No.: 62.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.21%  
DB: 9  
Gaps: 0  
  
US-09-513-151-3 (1-2041) x US-09-992-598-156 (1-378)  
OY 669 AGTCTCTAACCTTCATCCTT 692  
|||||  
Db 334 SerSerLeuthrLeuAlaSerPhe 341  
  
RESULT 8  
US-09-989-293A-156  
Sequence 156, Application US/09989293A  
Patent No. US2002017164A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Collin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OR INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C6  
CURRENT FILING DATE: 2001-11-20  
PRIOR APPLICATION NUMBER: US/09/989, 293A  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02

[illegible]

Alignment Scores:

Pred. NO.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-989-293A-156 (1-378)

OY 669 AGTCTCTAACCTTGATCCTT 692  
|||||

Db 334 SerSerLeuThrLeuAlaSerPhe 341

RESULT 9

US-09-989-735-156

Sequence 156, Application US/09989735

Publication No. US20020193299A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Bolstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Grimaldi, J. Christopher

APPLICANT: Guiney, Austin L.

APPLICANT: Kljavin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

TITLE OF INVENTION: Acids Encoding the Same

FILE REFERENCE: P2730P1C61

CURRENT APPLICATION NUMBER: US/09/989,735

CURRENT FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088025

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088028

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088029

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088030

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088326

PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088202

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088212

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088217

PRIOR FILING DATE: 1998-06-05

PRIOR APPLICATION NUMBER: 60/088655

PRIOR FILING DATE: 1998-06-09

PRIOR APPLICATION NUMBER: 60/088734

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088738

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088742

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088810

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088824

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858

PRIOR FILING DATE: 1998-06-11

PRIOR APPLICATION NUMBER: 60/088861

PRIOR FILING DATE: 1998-06-11

PRIOR APPLICATION NUMBER: 60/088876

PRIOR FILING DATE: 1998-06-11

PRIOR APPLICATION NUMBER: 60/089105

PRIOR FILING DATE: 1998-06-12

PRIOR APPLICATION NUMBER: 60/089440

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089512

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089514

PRIOR FILING DATE: 1998-06-16

PRIOR APPLICATION NUMBER: 60/089532

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089538

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089598

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089599

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089600

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089653

PRIOR FILING DATE: 1998-06-17

PRIOR APPLICATION NUMBER: 60/089801

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/089907

PRIOR FILING DATE: 1998-06-18

PRIOR APPLICATION NUMBER: 60/089908



[illegible]

```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.21%                Indels: 0
DB: 9                              Gaps: 0
US-09-513-151-3 (1-2041) x US-09-990-444-156 (1-378)

OY 669 AGTTCCTACCTGCATCCTT 692
DB 334 SerSerLeuThrLeuAlaSerPhe 341

RESULT 11
US-09-989-730-156
; Sequence 156, Application US/09989730
; Publication No. US20020197674A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerder, Hanspeter
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730PIC69
; CURRENT APPLICATION NUMBER: US/09/989,730
; PRIOR APPLICATION NUMBER: 2001-11-20
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948
```

PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091626  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

Alignment Scores:  
Pred. No.: 62.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%

Length: 378  
Matches: 8  
Conservative: 0  
Mismatches: 0

Query Match: 1.21% Indels: 0  
DB: 9 Gaps: 0  
US-09-513-151-3 (1-2041) x US-09-989-730-156 (1-378)  
Oy 669 AGTCTCTAACCTTGATCCTT 692  
|||||  
Db 334 SerSerLeuThrLeuAlaSerPhe 341  
RESULT 12  
US-09-990-436-156  
Sequence 156, Application US/09990436  
Publication No. US20020196148A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tunas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C14  
CURRENT APPLICATION NUMBER: US/09/990.436  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088025  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088026  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952

;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0



US-09-513-151-3 (1-2041) x US-09-990-436-156 (1-378)  
QY 669 AGTCTACCCCTTGATCTTT 692  
DB 334 SerSerLeuThrLeuAlaSerPhe 341  
RESULT 13  
US-09-991-181-156  
Sequence 156, Application US/09991181  
Publication No. US20020197615A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bolstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerltsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavir, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1C53  
CURRENT FILING DATE: 2001-11-16  
PRIOR APPLICATION NUMBER: US/09/991,181  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246

PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090431  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090435  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090444  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090535  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090540  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090542  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090676  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090678  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090690  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090694  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090696  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090862  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091478  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091544  
PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091636  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091633  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091978  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/092182  
PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.: 62.8 Length: 378  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.21% Indels: 0  
DB: 9 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-991-181-156 (1-378)

Qy 669 AGTCTCTAACCCCTTGACATCCTT 692  
Db 334 SerSerLeuThrLeuAlaSerPhe 341  
RESULT 14  
US-09-993-687-156  
Sequence 156, Application US/09993687  
Publication No. US20020198149A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Bostein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2730P1c11  
CURRENT APPLICATION NUMBER: US/09/993,687  
PRIOR FILING DATE: 2002-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04

;; PRIOR APPLICATION NUMBER: 60/088028  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088029  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088030  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088033  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088326  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/088167  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088202  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088212  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088217  
;; PRIOR FILING DATE: 1998-06-05  
;; PRIOR APPLICATION NUMBER: 60/088655  
;; PRIOR FILING DATE: 1998-06-09  
;; PRIOR APPLICATION NUMBER: 60/088734  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088738  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088742  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088810  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088824  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088826  
;; PRIOR FILING DATE: 1998-06-10  
;; PRIOR APPLICATION NUMBER: 60/088858  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088861  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/088876  
;; PRIOR FILING DATE: 1998-06-11  
;; PRIOR APPLICATION NUMBER: 60/089105  
;; PRIOR FILING DATE: 1998-06-12  
;; PRIOR APPLICATION NUMBER: 60/089440  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089512  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089514  
;; PRIOR FILING DATE: 1998-06-16  
;; PRIOR APPLICATION NUMBER: 60/089532  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089538  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089598  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089599  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089600  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089653  
;; PRIOR FILING DATE: 1998-06-17  
;; PRIOR APPLICATION NUMBER: 60/089801  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089907  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089908  
;; PRIOR FILING DATE: 1998-06-18  
;; PRIOR APPLICATION NUMBER: 60/089947  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252

;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-993-687-156 (1-378)

Oy 669 AGTCTCAACCTGCATCCTT 692

```
Db      334 SerSerLeuThrLeuAlaSerPhe 341
|||||
RESULT 15
US-09-989-734-156
: Sequence 156, Application US/09989734
: Publication No. US20030003531A1
GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi J.
: APPLICANT: Baker, Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Fong, Sherman
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Geritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Kijavlin, Ivar J.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K.
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides, and Nucleic
: FILE REFERENCE: P2730P1C64
: CURRENT APPLICATION NUMBER: US/09/989,734
: PRIOR FILING DATE: 2001-11-19
: PRIOR APPLICATION NUMBER: 60/049787
: PRIOR FILING DATE: 1997-06-16
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/065186
: PRIOR FILING DATE: 1997-11-12
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066770
: PRIOR FILING DATE: 1997-11-24
: PRIOR APPLICATION NUMBER: 60/075945
: PRIOR FILING DATE: 1998-02-25
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/083322
: PRIOR FILING DATE: 1998-04-28
: PRIOR APPLICATION NUMBER: 60/084600
: PRIOR FILING DATE: 1998-05-07
: PRIOR APPLICATION NUMBER: 60/087106
: PRIOR FILING DATE: 1998-05-28
: PRIOR APPLICATION NUMBER: 60/087607
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087609
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087759
: PRIOR FILING DATE: 1998-06-02
: PRIOR APPLICATION NUMBER: 60/087827
: PRIOR FILING DATE: 1998-06-03
: PRIOR APPLICATION NUMBER: 60/088021
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088025
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088026
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088028
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088029
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088030
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088033
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088326
: PRIOR FILING DATE: 1998-06-04
: PRIOR APPLICATION NUMBER: 60/088167
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088202
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088212
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088217
: PRIOR FILING DATE: 1998-06-05
: PRIOR APPLICATION NUMBER: 60/088555
: PRIOR FILING DATE: 1998-06-09
: PRIOR APPLICATION NUMBER: 60/088734
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088738
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088742
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088810
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088824
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088826
: PRIOR FILING DATE: 1998-06-10
: PRIOR APPLICATION NUMBER: 60/088858
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/088861
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/088876
: PRIOR FILING DATE: 1998-06-11
: PRIOR APPLICATION NUMBER: 60/089105
: PRIOR FILING DATE: 1998-06-12
: PRIOR APPLICATION NUMBER: 60/089440
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089512
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089514
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089532
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089538
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089598
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089599
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089600
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089653
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089801
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089907
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089908
: PRIOR FILING DATE: 1998-06-18
: PRIOR APPLICATION NUMBER: 60/089947
: PRIOR FILING DATE: 1998-06-19
: PRIOR APPLICATION NUMBER: 60/089948
: PRIOR FILING DATE: 1998-06-19
: PRIOR APPLICATION NUMBER: 60/089952
: PRIOR FILING DATE: 1998-06-19
: PRIOR APPLICATION NUMBER: 60/090246
: PRIOR FILING DATE: 1998-06-22
: PRIOR APPLICATION NUMBER: 60/090252
: PRIOR FILING DATE: 1998-06-22
: PRIOR APPLICATION NUMBER: 60/090254
```

Search completed: April 21, 2003, 19:34:14  
Job time : 41.5 secs

;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

## Alignment Scores:

Pred. No.:	62.8	Length:	378
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	9	Gaps:	0

US-09-513-151-3 (1-2041) x US-09-989-734-156 (1-378)

OY 669 AGTCTCTAACCTTGATCCTTT 692  
|||||  
DB 334 SerSerLeuThrLeuAlaSerPhe 341

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 21:27:17 : Search time 210 Seconds  
(without alignments)  
9763.385 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041  
Sequence: 1 CTGCCATAGATGCGCTCCG.....TTTACAGAGAAAAA 2041

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 639749 seqs, 502280978 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Published Applications.NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	473	23.2	717	10 US-09-764-853-166	Sequence 166, App
2	473	23.2	717	10 US-09-764-853-364	Sequence 364, App
3	318	15.6	318	9 US-09-764-692-4310	Sequence 4310, App
4	145	7.1	457	10 US-09-864-761-10050	Sequence 10050, A
5	138	6.8	138	10 US-09-864-761-26692	Sequence 26692, A
6	23	1.1	180557	12 US-10-003-806-6	Sequence 6, Appli
7	23	1.1	180557	12 US-10-003-806-9	Sequence 9, Appli
8	22	1.1	1118	10 US-09-764-877-2777	Sequence 2777, App
9	22	1.1	1398	10 US-09-764-877-2780	Sequence 2780, App
10	21	1.0	2154	9 US-09-822-846-301	Sequence 301, Appli
11	21	1.0	99916	10 US-09-812-095-3	Sequence 3, Appli
12	20	1.0	716	10 US-09-910-943-192	Sequence 192, App
13	20	1.0	1684	9 US-10-125-540-582	Sequence 582, App
14	20	1.0	1684	10 US-09-764-870-582	Sequence 582, App
15	20	1.0	1684	10 US-09-764-853-857	Sequence 857, App
16	20	1.0	1684	10 US-09-764-860-1035	Sequence 1035, App
17	20	1.0	1772	10 US-09-822-830A-214	Sequence 214, App
18	20	1.0	2396	9 US-10-125-540-581	Sequence 581, App
19	20	1.0	2396	10 US-09-764-870-581	Sequence 581, App

## ALIGNMENTS

C	20	20	1.0	2396	10	US-09-764-853-856	Sequence 856, App
C	21	20	1.0	2396	10	US-09-764-860-1034	Sequence 1034, App
C	22	20	1.0	11216	9	US-09-764-872-490	Sequence 490, App
C	23	20	1.0	69327	10	US-09-777-921A-3	Sequence 3, Appli
C	24	20	1.0	198285	10	US-09-880-107-3814	Sequence 3814, App
C	25	20	1.0	465237	10	US-09-933-267A-1	Sequence 1, Appli
C	26	19	0.9	213	9	US-10-083-357-206	Sequence 206, App
C	27	19	0.9	384	10	US-09-983-965-5442	Sequence 5442, App
C	28	19	0.9	594	9	US-10-002-344A-59	Sequence 59, Appli
C	29	19	0.9	1179	10	US-09-770-445-52	Sequence 52, Appli
C	30	19	0.9	1659	9	US-09-852-797-24	Sequence 24, Appli
C	31	19	0.9	1659	10	US-09-853-161-24	Sequence 24, Appli
C	32	19	0.9	1659	10	US-09-852-659A-24	Sequence 24, Appli
C	33	19	0.9	2848	9	US-10-002-344A-60	Sequence 60, Appli
C	34	19	0.9	3564	9	US-10-091-504-1986	Sequence 1986, App
C	35	19	0.9	3564	10	US-09-764-869-1986	Sequence 1986, App
C	36	19	0.9	4548	9	US-10-091-504-2034	Sequence 2034, App
C	37	19	0.9	4548	10	US-09-764-869-2034	Sequence 2034, App
C	38	19	0.9	12566	9	US-10-091-504-2035	Sequence 2035, App
C	39	19	0.9	12566	10	US-09-764-869-2035	Sequence 2035, App
C	40	19	0.9	18000	10	US-09-800-631-17	Sequence 17, Appli
C	41	19	0.9	42000	9	US-10-081-563-25	Sequence 25, Appli
C	42	19	0.9	49136	10	US-09-768-877-1	Sequence 1, Appli
C	43	19	0.9	180216	10	US-09-835-232-6	Sequence 6, Appli
C	44	19	0.9	397658	10	US-09-813-320-3	Sequence 3, Appli
C	45	18	0.9	41	10	US-09-766-273-15	Sequence 15, Appli

RESULT 1  
US-09-764-853-166  
Sequence 166, Application US/09764853  
Patent No. US2002009672A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P206  
CURRENT APPLICATION NUMBER: US/09/764, 853  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 166  
LENGTH: 717  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (619)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (625)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (655)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (673)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (713)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-853-166  
Query Match 23.2% Score 473; DB 10; Length 717;  
Best Local Similarity 99.7% Pred No. 3.3e-239;  
Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
44 GTGGCAGTGGCTCAGGGCTGCAACGACCTCTTGTAGTATTCGGGCC 103  
DB 44 GTGGCAGTGGCTCAGGGCTGCAACGACCTCTTGTAGTATTCGGGCC 103

```

Query Match: 23.28; Score: 473; DB: 10; Length: 717;
Best Local Similarity: 99.74; Pred. No.: 3; 3e-239;
Matches: 573; Conservative: 0; Mismatches: 2; Indels: 0; Gaps: 0;

QY 44 GTGGGCGAGTGGGCTTCAGGGGCGCTGCACAGGACCCCTACCTCTTGTAGTGATTCGCGGCC 103
DB 44 GTGGGCGAGTGGGCTTCAGGGGCGCTGCACAGGACCCCTACCTCTTGTAGTGATTCGCGGCC 103
QY 104 ACGGGACCGGGCAATATCCACGCTGGCGTTGAGCTAGGCCAGCGCGCTGGCGGGAGATC 163
DB 104 ACGGGACCGGGCAATATCCACGCTGGCGTTGAGCTAGGCCAGCGCGCTGGCGGGAGATC 163
QY 164 GTCAGCGCTGACCTCATGCGAGGCTATGAAAGCGCTAGACATCATCACCAAGTTTCT 223
DB 164 GTCAGCGCTGACCTCATGCGAGGCTATGAAAGCGCTAGACATCATCACCAAGTTTCT 223
QY 224 GCCCAAGGCGAGAAATTCGCGGCGCCACCATATATCATGCTTGTGGATCTCTTGTGACC 283
DB 224 GCCCAAGGCGAGAAATTCGCGGCGCCACCATATGATCATGCTTGTGGATCTCTTGTGACC 283
QY 284 AATTACACAGGCTGAGACTTTCAGAAATAGAGCACTGCTGATTTGAGATATATTTGGC 343
DB 284 AATTACACAGGCTGAGACTTTCAGAAATAGAGCACTGCTGATTTGAGATATATTTGGC 343
QY 344 CGAGACAAATTCCTATTTGTTGTGGAGGAACAAATTTATCATTTGATTCCTGCTCTGC 403
DB 344 CGAGACAAATTCCTATTTGTTGTGGAGGAACAAATTTATCATTTGATTCCTGCTCTGC 403
QY 404 AAAGTCTTGTCATPACCAAGCCCGAGAGATGGGCACTGGAAGGATTTGACCGGAAA 463
DB 404 AAAGTCTTGTCATPACCAAGCCCGAGAGATGGGCACTGGAAGGATTTGACCGGAAA 463
QY 464 GTGAGACTTGAAGAGAGATGATCTTGTACTTTCACAAAGCCCTAGCCAGGTGAGCCA 523
DB 464 GTGAGACTTGAAGAGAGATGATCTTGTACTTTCACAAAGCCCTAGCCAGGTGAGCCA 523
QY 524 GAAATGCTGCGCAAGCTGCATTCACATGACAAAGCGGCGAGAGCTTGCAAGTT 583
DB 524 GAAATGCTGCGCAAGCTGCATTCACATGACAAAGCGGCGAGAGCTTGCAAGTT 583
QY 584 TTGAGAAAGACGAGATCTCTCATAGTGAATTTCT 618
DB 584 TTGAGAAAGACGAGATCTCTCATAGTGAATTTCT 618

RESULT 3
US-09-796-692-4310
: Sequence 4310, Application US/09796692
: Publication No. US20020198362A1
: GENERAL INFORMATION:
: APPLICANT: Calger, Alexander
: APPLICANT: Algate, Paul A.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND T
: FILE REFERENCE: 2077_001200
: CURRENT APPLICATION NUMBER: US/09/796,692
: PRIOR APPLICATION NUMBER: 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: 60/200,599
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: 60/206,201

```

```

; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4310
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-4310

Query Match
Best Local Similarity 15.6%; Score 318; DB 9; Length 318;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1690 AAGAAGTTTATGTCCTGACTGCTGGCTAAATTAATTCAGATGCTTTGTAG 1749
DB 1 AAGAAGTTTATGTCCTGACTGCTGGCTAAATTAATTCAGATGCTTTGTAG 60
QY 1750 ATGACTGAGTATTTGTGACCAATTTGGAGTTCTAGATTTGAGTGAATGGCAGAA 1809
DB 61 ATGACTGAGTATTTGTGACCAATTTGGAGTTCTAGATTTGAGTGAATGGCAGAA 120
QY 1810 AGGCGCATCTCCATTTAGTGAATTAAGTGAACCAACTGTTCTCGAATTTCTACAGAGA 1869
DB 121 AGGCGCATCTCCATTTAGTGAATTAAGTGAACCAACTGTTCTCGAATTTCTACAGAGA 180
QY 1870 AGGAGGAATCAGACTGAGGAGAGCTGTGACATGAGACTTGAAGCAAGACTTTGAAT 1929
DB 181 AGGAGGAATCAGACTGAGGAGAGCTGTGACATGAGACTTGAAGCAAGACTTTGAAT 240
QY 1930 TTGCGAGCTGCTCATGTGAGATTATATCAGCTGCTCTTCTTATGACTTCAATCT 1989
DB 241 TTGCGAGCTGCTCATGTGAGATTATATCAGCTGCTCTTCTTATGACTTCAATCT 300
QY 1990 ATATTTTATTTAGATT 2007
DB 301 ATATTTTATTTAGATT 318

RESULT 4
US-09-864-761-10050/c
; Sequence 10050, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10050
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL033527.25
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-10050

Query Match
Best Local Similarity 7.1%; Score 145; DB 10; Length 457;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 AGGCTATGAGGCTTATGATCATCATCAGCAAGGTTTCTGCGCCAGAGAGATCT 242
DB 450 AGGCTATGAGGCTTATGATCATCATCAGCAAGGTTTCTGCGCCAGAGAGATCT 391
QY 243 GCCGCGCACCATGATCAGCTTTGTGATCTCTTGTGACCAATTACAGAGTGTGACT 302
DB 390 GCCGCGCACCATGATCAGCTTTGTGATCTCTTGTGACCAATTACAGAGTGTGACT 331
QY 303 TCAGAAATAGAGCACTGCTGAT 327
DB 330 TCAGAAATAGAGCACTGCTGAT 306

RESULT 5
US-09-864-761-26692/c
; Sequence 26692, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
```



Prior application data removed - refer to PALM or file wrapper  
 NUMBER OF SEQ ID NOS: 4031  
 SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2777  
LENGTH: 1118  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-877-2777

Query Match  
Best Local Similarity 100.0%; Score 22; DB 10; Length 1118;  
Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0;

QY 2019 AAATTTCACAGCAAAAAA 2040  
DB 221 AAATTTCACAGCAAAAAA 200

RESULT 9  
US-09-764-877-2780/C

Sequence 2780, Application US/09764877  
Patent No. US20020147140A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC005  
CURRENT APPLICATION NUMBER: US/09/764,877  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 4031  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 2780  
LENGTH: 1398  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-877-2780

Query Match  
Best Local Similarity 100.0%; Score 22; DB 10; Length 1398;  
Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0;

QY 2019 AAATTTCACAGCAAAAAA 2040  
DB 221 AAATTTCACAGCAAAAAA 200

RESULT 10  
US-09-822-846-301/C

Sequence 301, Application US/09822846  
Publication No. US20030027139A1  
GENERAL INFORMATION:  
APPLICANT: McCoy, John M.  
APPLICANT: Lavallie, Edward R.  
APPLICANT: Collins-Racie, Lisa A.  
APPLICANT: Evans, Cheryl  
APPLICANT: Merberg, David  
APPLICANT: Treacy, Maurice  
APPLICANT: Agostino, Michael J.  
APPLICANT: Steininger II, Robert J.  
APPLICANT: Bowman, Michael R.  
APPLICANT: Spaulding, Vikki  
APPLICANT: Wong, Gordon G.  
APPLICANT: Clark, Hilary  
APPLICANT: Fechtel, Kim  
APPLICANT: Howes, Steven H.  
APPLICANT: Resnick, Richard J.  
APPLICANT: Gulukota, Kamalakara  
APPLICANT: Graham, James R.  
APPLICANT: Genetics Institute, Inc.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
FILE REFERENCE: GIN 6400  
CURRENT APPLICATION NUMBER: US/09/822,846  
CURRENT FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/195,605  
PRIOR FILING DATE: 2000-04-06

NUMBER OF SEQ ID NOS: 629  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 301  
LENGTH: 2154  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-822-846-301

Query Match  
Best Local Similarity 100.0%; Score 21; DB 9; Length 2154;  
Pred. No. 1.7; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

QY 1566 CAGCAAGCATTTTTC 1586  
DB 1639 CAGCAAGCATTTTTC 1619

RESULT 11  
US-09-816-095-3

Sequence 3, Application US/09816095  
Patent No. US20020137164A1  
GENERAL INFORMATION:  
APPLICANT: Gan, Weiniu  
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
FILE REFERENCE: CLO01147  
CURRENT APPLICATION NUMBER: US/09/816,095  
CURRENT FILING DATE: 2001-03-26  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 99916  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(99916)  
OTHER INFORMATION: n = A,T,C or G  
US-09-816-095-3

Query Match  
Best Local Similarity 100.0%; Score 21; DB 10; Length 99916;  
Pred. No. 2.1; Mismatches 0; Indels 0; Gaps 0;  
Matches 21; Conservative 0;

QY 2019 AAATTTCACAGCAAAAAA 2039  
DB 43316 AAATTTCACAGCAAAAAA 43336

RESULT 12  
US-09-910-943-192

Sequence 192, Application US/09910943  
Patent No. US20020081610A1  
GENERAL INFORMATION:  
APPLICANT: Hemmati-Briandlou, Ali  
TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression  
FILE REFERENCE: 7529/1G1480S1  
CURRENT APPLICATION NUMBER: US/09/910,943  
CURRENT FILING DATE: 2001-07-23  
NUMBER OF SEQ ID NOS: 742  
SOFTWARE: Patentln version 3.1  
SEQ ID NO 192  
LENGTH: 716  
TYPE: DNA  
ORGANISM: Xenopus laevis  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(716)  
OTHER INFORMATION: n may be a or g or c or t/u  
US-09-910-943-192

Query Match 1.0%; Score 20; DB 10; Length 716;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2022 TTTACAGAGAAAAAAA 2041  
DB 492 TTTACAGAGAAAAAAA 511

## RESULT 13

US-10-125-540-582  
; Sequence 582, Application US/10125540  
; Publication No. US2003005985A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214C1  
; CURRENT APPLICATION NUMBER: US/10/125,540  
; CURRENT FILING DATE: 2002-04-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 582  
; LENGTH: 1684  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-125-540-582

Query Match 1.0%; Score 20; DB 9; Length 1684;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AAATTACAGAAAAAAA 2038  
DB 378 AAATTACAGAAAAAAA 397

## RESULT 14

US-09-764-870-582  
; Sequence 582, Application US/09764870  
; Patent No. US20020042386A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT214  
; CURRENT APPLICATION NUMBER: US/09/764,870  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 646  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 582  
; LENGTH: 1684  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-870-582

Query Match 1.0%; Score 20; DB 10; Length 1684;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AAATTACAGAAAAAAA 2038  
DB 378 AAATTACAGAAAAAAA 397

## RESULT 15

US-09-764-853-857/c  
; Sequence 857, Application US/09764853  
; Patent No. US20020090672A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PJ206

; CURRENT APPLICATION NUMBER: US/09/764,853  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 939  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 857  
; LENGTH: 1684  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-853-857

Query Match 1.0%; Score 20; DB 10; Length 1684;  
Best Local Similarity 100.0%; Pred. No. 5.8;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AAATTACAGAAAAAAA 2038  
DB 1307 AAATTACAGAAAAAAA 1288

Search completed: April 22, 2003, 00:12:53  
Job time : 678 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 21:22:03 ; Search time 968 Seconds

(Without alignments)  
9452.049 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 2041  
Sequence: 1 CTCGCATTAACATGCGTCGCG.....TTTACAGCAAAAAA 2041

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 5930009 seqs, 224146263 residues

Word size: 0

Total number of hits satisfying chosen parameters: 11860018

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

1: Pending-Patents\_NA\_New:\*  
2: /cgn2\_6/pdata/2/pna/US06\_NEW.COMB.seq:\*  
3: /cgn2\_6/pdata/2/pna/US07\_NEW.COMB.seq:\*  
4: /cgn2\_6/pdata/2/pna/US08\_NEW.COMB.seq:\*  
5: /cgn2\_6/pdata/2/pna/US09\_NEW.COMB.seq:\*  
6: /cgn2\_6/pdata/2/pna/US09\_NEW.COMB.seq2:\*  
7: /cgn2\_6/pdata/2/pna/US09\_NEW.COMB.seq3:\*  
8: /cgn2\_6/pdata/2/pna/US10\_NEW.COMB.seq:\*  
9: /cgn2\_6/pdata/2/pna/US10\_NEW.COMB.seq2:\*  
10: /cgn2\_6/pdata/2/pna/US60\_NEW.COMB.seq:\*  
11: /cgn2\_6/pdata/2/pna/US60\_NEW.COMB.seq2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Query	Match	Length	DB	ID	Description
No.	Score					
1	896	43.9	2130	6	US-09-659-671A-134	Sequence 134, App
2	896	43.9	2130	8	US-10-380-731-119	Sequence 119, App
3	767	37.6	1271	8	US-10-170-235-41124	Sequence 41124, A
4	765	37.5	1719	8	US-10-170-235-41123	Sequence 41123, A
5	765	37.5	2119	8	US-10-170-235-40554	Sequence 40554, A
6	751	36.8	1749	9	US-10-342-887-1692	Sequence 1692, App
7	622	30.5	1271	11	US-60-453-135-276	Sequence 276, App
8	622	30.5	1271	11	US-60-453-050-276	Sequence 276, App
9	622	30.5	1719	11	US-60-453-135-274	Sequence 274, App
10	622	30.5	1719	11	US-60-453-050-274	Sequence 274, App
11	622	30.5	2119	11	US-60-453-135-275	Sequence 275, App
12	622	30.5	2119	11	US-60-453-050-275	Sequence 275, App
13	404	19.8	555	6	US-09-532-315B-3472	Sequence 3472, App
14	269	13.2	392	6	US-09-532-315B-3471	Sequence 3471, App
15	265	13.0	293	6	US-09-532-315B-9521	Sequence 9521, App
16	264	12.9	264	6	US-09-532-315B-9514	Sequence 9514, App
17	256	12.5	255	6	US-09-532-315B-9508	Sequence 9508, App
18	255	12.5	255	6	US-09-532-315B-9511	Sequence 9511, App
19	247	12.1	247	6	US-09-532-315B-9505	Sequence 9505, App
20	237	11.6	266	6	US-09-532-315B-9507	Sequence 9507, App
21	230	11.3	289	6	US-09-532-315B-9519	Sequence 9519, App
22	193	9.5	240	6	US-09-532-315B-9515	Sequence 9515, App

23	192	9.4	265	6	US-09-532-315B-9518	Sequence 9518, App
24	189	9.3	189	6	US-09-532-315B-9500	Sequence 9500, App
25	182	8.9	245	6	US-09-532-315B-9523	Sequence 9523, App
26	180	8.8	263	6	US-09-532-315B-9510	Sequence 9510, App
27	179	8.8	180	6	US-09-532-315B-9503	Sequence 9503, App
28	174	8.5	334	6	US-09-912-293-109604	Sequence 109604, App
29	172	8.4	232	6	US-09-532-315B-9529	Sequence 9529, App
30	150	7.3	201	11	US-60-453-135-18354	Sequence 18354, A
31	150	7.3	201	11	US-60-453-135-18357	Sequence 18357, A
32	150	7.3	201	11	US-60-453-135-18362	Sequence 18362, A
33	150	7.3	201	11	US-60-453-135-18364	Sequence 18364, A
34	150	7.3	201	11	US-60-453-135-18368	Sequence 18368, A
35	150	7.3	201	11	US-60-453-050-18354	Sequence 18354, A
36	150	7.3	201	11	US-60-453-050-18357	Sequence 18357, A
37	150	7.3	201	11	US-60-453-050-18362	Sequence 18362, A
38	150	7.3	201	11	US-60-453-050-18364	Sequence 18364, A
39	150	7.3	201	11	US-60-453-050-18368	Sequence 18368, A
40	149	7.3	200	6	US-09-532-315B-9517	Sequence 9517, App
41	148	7.3	255	6	US-09-532-315B-9520	Sequence 9520, App
42	145	7.1	457	9	US-10-203-138A-1534	Sequence 1534, App
43	144	7.1	262	6	US-09-532-315B-9528	Sequence 9528, App
44	139	6.8	500	6	US-09-912-293-111181	Sequence 111181, App
45	138	6.8	138	9	US-10-203-138A-6661	Sequence 6661, App

## ALIGNMENTS

```
RESULT 1
US-09-659-671A-134
Sequence 134, Application US/09659671A
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chennhua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyun
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wehtman, Tom
APPLICANT: Drmanac, Radolje T.
TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 794
CURRENT APPLICATION NUMBER: US/09/659, 671A
CURRENT FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 475
SOFTWARE: pf_fl_genes Version 2.0
SEQ ID NO 134
LENGTH: 2130
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (191)..(1426)
US-09-659-671A-134
Query Match 43.9%; Score 896; DB 6; Length 2130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 44 GTGGCAGTGGCTAGAGGCGCTGACAGGACCTACTTGTAGTATTTCGGGGCC 103  
|||||  
DB 56 GTGGCAGTGGCTAGAGGCGCTGACAGGACCTACTTGTAGTATTTCGGGGCC 115  
|||||  
QY 104 ACGGCGACCGGCAATTCACGCTGCTAGGCGACGCGCGCTGAGATC 163  
|||||  
DB 116 ACGGCGACCGGCAATTCACGCTGCTAGGCGACGCGCGCTGAGATC 175  
|||||  
QY 164 GTGACGCTGACTCCATGCTATGAGAGGCTATGACCAACAGAGTCTTCT 223

```
|||||
Db 176 GTGAGCCCTGACTCCATCGACAGTGTCTATGAAGCCCTACACATCATCACACAAAGTTTCT 235
OY 224 GCCCAGAGCAGAGAAATCTCCGGCCACACATGATGAGCTTTGTGATCTCTTGATCC 283
Db 236 GCCCAAGAGCAGAGAAATCTCCGGCCACACATGATGAGCTTTGTGATCTCTTGATCC 295
OY 284 AATTACACAGTGGTGGGACTTCAGAAATAGAGCAAACTGCTGTGATGGAATATATTGGCC 343
Db 296 AATTACACAGTGGTGGGACTTCAGAAATAGAGCAAACTGCTGTGATGGAATATATTGGCC 355
OY 344 CGAGACAAATATCTTATTTGTGGGAGAGCAATATTATATGATGATGATCTGCTGAG 403
Db 356 CGAGACAAATATCTTATTTGTGGGAGAGCAATATTATATGATGATGATCTGCTGAG 415
OY 404 AAAGTTTGTCTCAATACCAAGCCCGCAGAGATGGGCACTGAGAAATGATTTGACCGAAA 463
Db 416 AAAGTTTGTCTCAATACCAAGCCCGCAGAGATGGGCACTGAGAAATGATTTGACCGAAA 475
OY 464 GTGAGCTTGAAGAGAGATGGTCTGTACTTCCACAAAGCCCTAACCGAGTGGAGCCA 523
Db 476 GTGAGCTTGAAGAGAGATGGTCTGTACTTCCACAAAGCCCTAACCGAGTGGAGCCA 535
OY 524 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAAGTT 583
Db 536 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAAGTT 595
OY 584 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 643
Db 596 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 655
OY 644 GGTGGTCCCTTGAGAGTCCCTGGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 703
Db 656 GGTGGTCCCTTGAGAGTCCCTGGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 715
OY 704 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGCTGCTGG 763
Db 716 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGCTGCTGG 775
OY 764 CTCTTGAGAGAACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 823
Db 776 CTCTTGAGAGAACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 835
OY 824 AGCAGAGCTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTTCAGCATAC 883
Db 836 AGCAGAGCTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTTCAGCATAC 895
OY 884 CTGATCACTGAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 939
Db 896 CTGATCACTGAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 951

RESULT 2
US-10-380-731-119
; Sequence 119, Application US/10380731
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-114
; CURRENT APPLICATION NUMBER: US/10/380,731
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/659,671
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 888
; SOFTWARE: Custom
; SEQ ID NO 119
; LENGTH: 2130
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (191)..(1426)
US-10-380-731-119
```

```
Query Match 43.9%; Score 896; DB 8; Length 2130;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 44 GTGGGCGAGTGGGCTCAGGGCCCTGCAACGAGACCTTACTCTTGTAGTATCTCGGGGCC 103
Db 56 GTGGGCGAGTGGGCTCAGGGCCCTGCAACGAGACCTTACTCTTGTAGTATCTCGGGGCC 115
OY 104 ACGGGCCACGGCAATTCACAGCTGCGCTTGCCAGCTAGGCGCAGCGGCTGGCGGTGAGATC 163
Db 116 ACGGGCCACGGCAATTCACAGCTGCGCTTGCCAGCTAGGCGCAGCGGCTGGCGGTGAGATC 175
OY 164 GTCAAGCCTACTCTCATGACAGTCTATGAAGGCTTAGACATCATCACAACAAAGTTTCT 223
Db 176 GTCAAGCCTACTCTCATGACAGTCTATGAAGGCTTAGACATCATCACAACAAAGTTTCT 235
OY 224 GCCCAAGAGCAGAGAAATCTCCGGCCACACATGATGAGCTTTGTGATCTCTTGATCC 283
Db 236 GCCCAAGAGCAGAGAAATCTCCGGCCACACATGATGAGCTTTGTGATCTCTTGATCC 295
OY 284 AATTACACAGTGGTGGGACTTCAGAAATAGAGCAAACTGCTGTGATGGAATATATTGGCC 343
Db 296 AATTACACAGTGGTGGGACTTCAGAAATAGAGCAAACTGCTGTGATGGAATATATTGGCC 355
OY 344 CGAGACAAATATCTTATTTGTGGGAGAGCAATATTATATGATGATGATCTGCTGAG 403
Db 356 CGAGACAAATATCTTATTTGTGGGAGAGCAATATTATATGATGATGATCTGCTGAG 415
OY 404 AAAGTTTGTCTCAATACCAAGCCCGCAGAGATGGGCACTGAGAAATGATTTGACCGAAA 463
Db 416 AAAGTTTGTCTCAATACCAAGCCCGCAGAGATGGGCACTGAGAAATGATTTGACCGAAA 475
OY 464 GTGAGCTTGAAGAGAGATGGTCTGTACTTCCACAAAGCCCTAACCGAGTGGAGCCA 523
Db 476 GTGAGCTTGAAGAGAGATGGTCTGTACTTCCACAAAGCCCTAACCGAGTGGAGCCA 535
OY 524 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAAGTT 583
Db 536 GAAATGGCTGCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAAGTT 595
OY 584 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 643
Db 596 TTGGAAGAAACAGCAATCTCTCATAGTGAATTTCTCATGTCACATACGGAAGAAGT 655
OY 644 GGTGGTCCCTTGAGAGTCCCTGGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 703
Db 656 GGTGGTCCCTTGAGAGTCCCTGGAAGTCTCTCAACCCCTTGACCTTTGGCTTCATGCT 715
OY 704 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGCTGCTGG 763
Db 716 GACCAGCAGTCTAGATGAGCGCTTGATAGAGGCTGATGACATGCTTGCTGCTGG 775
OY 764 CTCTTGAGAGAACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 823
Db 776 CTCTTGAGAGAACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAAT 835
OY 824 AGCAGAGCTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTTCAGCATAC 883
Db 836 AGCAGAGCTATCAACATGATCTTCCATCAATGAGCTTCAAGGAATTTTCAGCATAC 895
OY 884 CTGATCACTGAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 939
Db 896 CTGATCACTGAGGAAATGCACTGTGAGACTAGTAACCAAGCTTCTAAGAAAG 951

RESULT 3
US-10-170-235-41124
; Sequence 41124, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF
; TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREO
```

```

: FILE REFERENCE: CLO001380
: CURRENT APPLICATION NUMBER: US/10/170,235
: CURRENT FILING DATE: 2003-03-17
: NUMBER OF SEQ ID NOS: 42514
: SEQ ID NO 41124
: LENGTH: 1271
: TYPE: DNA
: ORGANISM: HUMAN
: US-10-170-235-41124

```

Query Match	37.68;	Score 767;	DB 8;	Length 1271;
Best Local Similarity	99.58;	Pred. No. 0;		
Matches 1087;	Conservative	0;	Mismatches 4;	Indels 1;
			Gaps	1;

Oy	937	AGGCCGTCGTCATTTGTCCTCCCTCTGATGGCTTAAAGGATATGATGATGTCGTGAAGTC	996
Db	181	AGGACCTGGTCCATTTGTCCCTCTCTATGGCTTTAGGGATATGTGATGTCGTGAAGTC	240
Oy	997	GGAGAGCTGTCTTTGGAACCTGCTCTTGAAAATCGTCAAAGTTTCATCCAGGGCCACAA	1056
Db	241	GGAGAGCTGTCTTTGGAACCTGCTCTTGAAAATCGTCAAAGTTTCATCCAGGGCCACAA	300
Oy	1057	GCTTACAGCCACATCTCCAAATGAAGTGGCCATTCATATGAAGTGAAGCAGAAGATTTATCA	1116
Db	301	GCCATCAGCCACATCTCCAAATGAAGTGGCCATTCATATGAAGTGAAGCAGAAGATTTATCA	360
Oy	1117	CCCTGTGACCTCTGTGATGATTCATTCATTTGGGGATCCGAAATGGGACGGACATATA	1176
Db	361	CCCTGTGACCTCTGTGATGATTCATTCATTTGGGGATCCGAAATGGGACGGACATATA	420
Oy	1177	ATCCAAATCCCACTTGAACCACTGAAGAAAAGAAAGAGATTGGACCTGATGCTGTCAA	1236
Db	421	ATCCAAATCCCACTTGAACCACTGAAGAAAAGAAAGAGATTGGACCTGATGCTGTCAA	480
Oy	1237	CACCATGAAAAGTAGAGTCTTCCGACACTATTAACAAGAACCTTAAGAAAGAGAGATC	1296
Db	481	CACCATGAAAAGTAGAGTCTTCCGACACTATTAACAAGAACCTTAAGAAAGAGAGATC	540
Oy	1297	CCGAGGCGAGATGATCAAGAGCTGGAATGCAAGCTTTAAGAGACTGTCCAGTGGCCTT	1356
Db	541	CCGAGGCGAGATGATCAAGAGCTGGAATGCAAGCTTTAAGAGAGACTGTCCAGTGGCCTT	600
Oy	1357	TGGAAAGCTGTGGGGATCCAGTTCAAGAGAGAGAGGGGTATGTTGTCCTCCACAGTGTGGC	1416
Db	601	TGGAAAGCTGTGGGGATCCAGTTCAAGAGAGAGAGGGGTATGTTGTCCTCCACAGTGTGGC	660
Oy	1417	AAAGAGATGCTATCCGAATTCCTCTCATAGCAGAAACCTCCACAAATTTCTTGTAT	1476
Db	661	AAAGAGATGCTATCCGAATTCCTCTCATAGCAGAAACCTCCACAAATTTCTTGTAT	720
Oy	1477	GTGCTTTTAAAGTCTCAGCTTCTCTATATATAGAAACAGAGCTCTTGTACGTCCTGTG	1536
Db	721	GTGCTTTTAAAGTCTCAGCTTCTCTATATATAGAAACAGAGCTCTTGTACGTCCTGTG	780
Oy	1537	TGGTGATGTCGCGAAATGATGATGTCAGAAAGCAATTTTCTTGTGAACCTT	1596
Db	781	TGGTGATGTCGTCGAAATGATGATGTCAGAAAGCAATTTTCTTGTGAACCTT	840
Oy	1597	AAAGTTCATATTTAAAGCAGACAGATTCACATTTTATACATGAGATCTCTTT	1656
Db	841	AAAGTTCATATTTAAAGCAGACAGATTCACATTTTATACATGAGATCTCTTT	900
Oy	1657	GTGCTGAATACGAGATTTGACTGCATCCCTTTAAAGAAGTTTATGTCCTCGAATCGG	1716
Db	901	GTGCTGAATACGAGATTTGACTGCATCCCTTTAAAGAAGTTTATGTCCTCGAATCGG	959
Oy	1717	CTAAATATATCTAATTTCCAGATGCTTTGTATGATACAGAGATATTTGTGACACATA	1776
Db	960	CTAAATATATCTAATTTCCAGATGCTTTGTATGATACAGAGATATTTGTGACACATA	1019
Oy	1777	TTGGAGCTTCAGATTTTGAGTGAATGGCAGAAAGGCCATCTCCATTTGATGATTTAAG	1836
Db	1020	TTGGAGCTTCAGATTTTGAGTGAATGGCAGAAAGGCCATCTCCATTTGATGATTTAAG	1079

Accession	Sequence	Length
Oy	TGACCAAACTAGTTCGGCAATTCACAGAGAGAGGAACTACAGCTGAGGAAAGCGT	1896
Db	TGAACCAAACTAGTTCGCAAAATTCACAGAGAGAGGAACTACAGCTGAGGAAAGCGT	1139
Oy	GACATAGGACTTGAAGACCAAAAGACTTGAATTTGCGAGCGTCATCTGTGACTATT	1956
Db	GACATAGGACTTGAAGACCAAAAGACTTGAATTTGCGAGCGTCATCTGTGACTATT	1199
Oy	ATCACTGCTGCTTCTATTGAGTTCACAAATCTATATTTTATTGAGTTTAAATAAAGA	2016
Db	ATCACTGCTGCTTCTATTGAGTTCACAAATCTATATTTTATTGAGTTTAAATAAAGA	1259
Oy	AAAAATTACAA 2028	
Db	AAAAATTACAA 1271	

```

RESULT 4
US-10-170-235-41123
; Sequence 41123, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THERE
; FILE REFERENCE: C1001380
; CURRENT APPLICATION NUMBER: US/10/170.235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 41123
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-41123

Query Match
Best Local Similarity 37.5%; Score 765; DB 8; Length 1719;
Matches 1085; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 939 GACCTGGTCCATTTGCCCCCTGTCATATGCTTAGAGATATCATATGCTTGAAGTGGG 998
DB 631 GACCTGGTCCATTTGCCCCCTGTCATATGCTTAGAGATATCATATGCTTGAAGTGGG 990
QY 999 AAGAGCTGTTCTTGTGACCTGCTCTTGAATGCTGAAGAATTCTCATCCAGGSCCAAGC 1058
DB 691 AAGAGCTGTTCTTGTGACCTGCTCTTGAATGCTGAAGAATTCTCATCCAGGSCCAAGC 750
QY 1059 CTACAGCCACTCCATAAAGATGCCATACAAATGAAGCTGAGAACAGAAAGATTATCACC 1118
DB 751 CTACAGCCACTCCATAAAGATGCCATACAAATGAAGCTGAGAACAGAAAGATTATCACC 810
QY 1119 TGTGTGACCTCTGTGATGCAATCATCATTTGGGGATTCGCAATGGGCGGCACTAAAT 1178
DB 811 TGTGTGACCTCTGTGATGCAATCATCATTTGGGGATTCGCAATGGGCGGCACTAAAT 870
QY 1179 CCAAAATCCCACTTGAACCACTGAGAAGAAAGAAAGATTGGAGCTCAGATGCTGTCAACA 1238
DB 871 CCAAAATCCCACTTGAACCACTGAGAAGAAAGAAAGATTGGAGCTCAGATGCTGTCAACA 930
QY 1239 CCATGAAAGTCAGAGTGTTCGCCAGACTATACAAAGAACCTAAAGGGAAGATCCC 1298
DB 931 CCATGAAAGTCAGAGTGTTCGCCAGACTATACAAAGAACCTAAAGGGAAGATCCC 990
QY 1299 CAGGGCAGAAATCATCAAGGCTGAATGAGGCTTAAAGAGCATGTCAGAGGGCTTTG 1358
DB 991 CAGGGCAGAAATCATCAAGGCTGAATGAGGCTTAAAGAGCATGTCAGAGGGCTTTG 1050
QY 1359 GAAAGGTGGTGGGAGATCCAGTTCAGAGAGGAGAGGGGTATGTTGTTCCAGTCTGGCAA 1418
DB 1051 GAAAGGTGGTGGGAGATCCAGTTCAGAGAGGAGAGGGGTATGTTGTTCCAGTCTGGCAA 1110
QY 1419 AGGAGTGTCTATGCGAATTTCTGCAATAGCAGAAAAGCTCCACCATTTTCTTTGATGT 1478

```

Db	1111	AGGAGTCTATGCGGAATTCCTCTGCATGACGAAAAAGCTCCACCATTCTTTTGATGT	1170
Qy	1479	GGTTTAAAGTCCACGTTCTCTATATATAGAAGCAGCGTCTGTGACCTGTGTG	1538
Db	1171	GGTTTAAAGTCCACGTTCTCTATATATAGAAGCAGCGTCTGTGACCTGTGTG	1230
Qy	1539	GCTGATGCTGTGCAAAATGATGTAGTTCAGGAAAGCATTTTTTTTCTTTGACCTTAA	1598
Db	1231	GCTGATGCTGTGCAAAATGATGTAGTTCAGGAAAGCATTTTTTTTCTTTGACCTTAA	1290
Qy	1599	AGGTTCTATTAATTAAGCAGCAGATTCACATTTTATATACATGAGATCTTCTGTG	1658
Db	1291	AGGTTCTATTAATTAAGCAGCAGATTCACATTTTATATACATGAGATCTTCTGTG	1350
Qy	1659	GGTGAATACACGAGATGTGACTGCATCCCTTTTAAAGAACTTTATGTCCCTGACTGTGCT	1718
Db	1351	GGTGAATACACGAGATGTGACTGCATCCCTTTTAAAGAACTTTATGTCCCTGACTGTGCT	1409
Qy	1719	AAATTTACTTAATTTCCGATGCTTTTGTACATGACTAAATATTTTGAGCCACATATT	1778
Db	1410	AAATTTACTTAATTTCCGATGCTTTTGTACATGACTAAATATTTTGAGCCACATATT	1469
Qy	1779	GGGAGTCTTAATTTGACTGAATGGCAGAAAGGCCATCTCCATTGAGATGATTAATG	1838
Db	1470	GGGAGTCTTAATTTGACTGAATGGCAGAAAGGCCATCTCCATTGAGATGATTAATG	1529
Qy	1839	AACCAACTAGTCTCGGAATTTACAGAGAAGGAGGAATCAGACTAGGAAAGCTGGA	1898
Db	1530	AACCAACTAGTCTCGGAATTTACAGAGAAGGAGGAATCAGACTAGGAAAGCTGGA	1589
Qy	1899	CATGAGACTTGAACCAACCAAGACTTTGAAATTTTGGACCTGCTCATGTGTAGTTATAT	1958
Db	1590	CATGAGACTTGAACCAACCAAGACTTTGAAATTTTGGACCTGCTCATGTGTAGTTATAT	1649
Qy	1959	CACGCTGCTCTTTCTATATGAGATTACAATCTATATTTTATTTGAAGTTAAATAAGAAA	2018
Db	1650	CACGCTGCTCTTTCTATATGAGATTACAATCTATATTTTATTTGAAGTTAAATAAGAAA	1709
Qy	2019	AAATTTACAA 2028	
Db	1710	AAATTTACAA 1719	

```

RESULT 5
US-10-170-235-40554
; Sequence 40554, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: C1001380
; CURRENT APPLICATION NUMBER: US/10/170.235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 40554
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-40554

Query Match          37.5%; Score 765; DB 8; Length 2119;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1085; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

OY 939 GACCTGGCCCATGTCCTCCCTGCTATGCGCTTAGAGGATATCGATGTCGCAAGTGG 998
      |||||||
Db 1031 GACCTGGTCCCATGTCCTCCCTGCTATGCGCTTAGAGGATATCGATGTCGCAAGTGG 1090

OY 999 AGGAGTCTGTTCTTGAACCTGCTCTGAAATCGTGCAAGTTTCATCCAGGGCCACAAGC 1058
      |||||||
Db 1091 AAGAGTCTGTTCTTGAACCTGCTCTGAAATCGTGCAAGTTTCATCCAGGGCCACAAGC 1150

OY 1059 CTACAGCCACCTCCATAAAGATGCCATACATGAAGCTGAGACAGAGAAGATTATCACC 1118

```

Dd	1151	CTACAGCCACTCCATTAAGATGTCCTATACATATGAACCTTGAGAACAGAAAGATTATACCC	1210
Oy	1119	TGTGTGACCTTGTGATCGAATCATCATTTGSGGATCGCAATGGCAGCGCACATAAAT	1178
Dd	1211	TGTGTGACCTCTGTGATCGAATCATCATTTGSGGATCGCAATGGCAGCGCACATAAAT	1270
Oy	1179	CCAAATCCCACTTGACCACTGACAGAAAAGAAAGAAATGGACTGATGCTGTCAACA	1238
Dd	1271	CCAAATCCCACTTGACCACTGACAGAAAAGAAAGAAATGGACTGATGCTGTCAACA	1330
Oy	1239	CCATAGAAAGCACAAGTGTTCCTCCAGACATATACAAACACTTAAGGAAAGGATATCC	1298
Dd	1331	CCATAGAAAGCACAAGTGTTCCTCCAGACATATACAAACACTTAAGGAAAGGATATCC	1390
Oy	1299	CAGGCGAGAATGATCAAGACCTGAATTCGACGCTTTAAGAGACATGTCAGTGGCTTTG	1358
Dd	1391	CAGGCGAGAATGATCAAGACCTGAATTCGACGCTTTAAGAGACATGTCAGTGGCTTTG	1450
Oy	1339	GAAAGGTGTGTGGGATTCACACTTGAGAGAGGAGGGGTATGTTGTCTCCAGTGTGGCAA	1418
Dd	1431	GAAAGGTGTGTGGGATTCACACTTGAGAGAGGAGGGGTATGTTGTCTCCAGTGTGGCAA	1510
Oy	1419	AGGAGTGTCTATGCGGATTTCTGCAATAGACAGAAAGCTCCACACATTTCTTTGATGT	1478
Dd	1511	AGGAGTGTCTATGCGGATTTCTGCAATAGACAGAAAGCTCCACACATTTCTTTGATGT	1570
Oy	1479	GCTTTTAAAGTCTCAGCTTCTCTATATATAGAAACACAGAGCTGTGTCAAGCTCTTGTTG	1538
Dd	1571	GCTTTTAAAGTCTCAGCTTCTCTATATATAGAAACACAGAGCTGTGTCAAGCTCTTGTTG	1630
Oy	1539	GCTGATGTGCTGGAAATGATGTATGTATGAGAAAGACTTTTCTTTTGTGAACCTTAA	1598
Dd	1631	GCTGATGTGCTGGAAATGATGTATGTATGAGAAAGACTTTTCTTTTGTGAACCTTAA	1690
Oy	1599	AGGTTCTATTATTTAAAGACGACAGACTTCCACATTTTATACATGAGGATCTTCTTGT	1658
Dd	1691	AGGTTCTATTATTTAAAGACGACAGACTTCCACATTTTATACATGAGGATCTTCTTGT	1750
Oy	1659	GGTAATATACAGAGTTACAGCATCCCTTTAAAGAGTTTATATGCCCTGACCTGGCT	1718
Dd	1751	GGTAATATACAGAGTTACAGCATCCCTTTAAAGAGTTTATATGCCCTGACCTGGCT	1809
Oy	1719	AAAATTAATCTAATTTCCAGATGCTTTTGTAGATGACTGAGATTTGTGAGCCACATATT	1778
Dd	1810	AAAATTAATCTAATTTCCAGATGCTTTTGTAGATGACTGAGATTTGTGAGCCACATATT	1869
Oy	1779	GGGAGTTCTAGATTTGAGTAATGCGACGAGAAAGGCCATCTCCATTGAGATGATTAAGT	1838
Dd	1870	GGGAGTTCTAGATTTGAGTAATGCGACGAGAAAGGCCATCTCCATTGAGATGATTAAGT	1929
Oy	1839	AACCAAACTAGTTCTCGGATTTCTACAGAGAAAGAGGAGANTCAGACTGAGAAAGCTGTGA	1898
Dd	1930	AACCAAACTAGTTCTCGGATTTCTACAGAGAAAGAGGAGANTCAGACTGAGAAAGCTGTGA	1989
Oy	1899	CATAGGACCTGGAAGACCAGAAAGCTTTGAAATTTGCGAGCTGCTCATGTGTGAGTTATAT	1958
Dd	1990	CATAGGACCTGGAAGACCAGAAAGCTTTGAAATTTGCGAGCTGCTCATGTGTGAGTTATAT	2049
Oy	1959	CACGCTGCTCTTCTATTTGATGTACAAATCTATTTTATTTATGAAGTTTAAATTAAGAAA	2018
Dd	2050	CACGCTGCTCTTCTATTTGATGTACAAATCTATTTTATTTATGAAGTTTAAATTAAGAAA	2109
Oy	2019	AAATTTACAA 2028	
Dd	2110	AAATTTACAA 2119	

```

RESULT 6
US-10-342-887-1692
; Sequence 1692, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue

```





	Query Match	30.5%	Score 622	DB 11:	Length 1271:
	Best Local Similarity	99.4%:	Pred.	No. 1.4e-298:	
	Matches	992:	Conservative	0:	Mismatches 5; Indels 1;
Oy	1031	GTCCAAAGTTTCATCTCAGGGCCACCAAGCTTACAGCCACTCCATTATGAAGTGCCATTACAAT	1090		
bcb	275	GTCGAAAATTTTCATCTCAGGGCCACCAAGCTTACAGCCACTCCATTATGAAGTGCCATTACAAT	334		

RESULT 9 -  
US-60-453-135-274  
Sequence 274, Application US/60453135  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: IAKOUBOVA, Olga  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

```

: TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001456
: CURRENT APPLICATION NUMBER: US/60/453,135
: CURRENT FILING DATE: 2003-03-10
: NUMBER OF SEQ ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 274
: LENGTH: 1719
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-60-453-135-274

```

```

Query Match          30.5%; Score 622; DB 11; Length 1719;
Best Local Similarity 99.4%; Pred. No. 1.4e-298;
Matches 992; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1031 GTGCAAGATTTCATCCAGGGCCACAAAGCCTACAGCCACTCCCAATTAAGTGCCTATCAAT 1090
DB 723 GTGCAAAAGTTTCATCCAGGGCCACAAAGCCTACAGCCACTCCCAATTAAGTGCCTATCAAT 782
QY 1091 GAAGCTGAGCAAGAGAGTATACCTGTGACCTGTGATGATGATGATGATGATGATGATGAT 1150
DB 783 GAAGCTGAGCAAGAGAGTATACCTGTGACCTGTGATGATGATGATGATGATGATGATGATGAT 842
QY 1151 GATCGGAGTGGGACGCGACATTAATAATCCCAATCCCACTTAACCACTGAAGAAAGA 1210
DB 843 GATCGGAGTGGGACGCGACATTAATAATCCCAATCCCACTTAACCACTGAAGAAAGA 902
QY 1211 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCCTCCAGACTAT 1270
DB 903 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCCTCCAGACTAT 962
QY 1271 AACCAAGAACTTAAGGAGAGGATCCCAAGGCGAGATGATCAAGAGCTGAATATCCAGC 1330
DB 963 AACCAAGAACTTAAGGAGAGGATCCCAAGGCGAGATGATCAAGAGCTGAATATCCAGC 1022
QY 1331 GTTTAAGACATGTCACATGGGCTTTGGAAGAGTGTGGGATCCAGTTCCAGAGAGAG 1390
DB 1023 GTTTAAGACATGTCACATGGGCTTTGGAAGAGTGTGGGATCCAGTTCCAGAGAGAG 1082
QY 1391 GGGTATGTTGTCCTCCACATCGGCGCAAGAGTGTGCGAATTTCTGTGCATAGCAG 1450
DB 1083 GGGTATGTTGTCCTCCACATCGGCGCAAGAGTGTGCGAATTTCTGTGCATAGCAG 1142
QY 1451 AAAAGCTCCACCATTTTCTTTGATGCTTTTAAGTCTACAGTTCTCTATATAGAA 1510
DB 1143 AAAAGCTCCACCATTTTCTTTGATGCTTTTAAGTCTACAGTTCTCTATATAGAA 1202
QY 1511 ACAGCAGTCTGTGACGTCCTGTTGCTGTGCTGATGCTGGAATAATGATGATGATGATGAT 1570
DB 1203 ACAGCAGTCTGTGACGTCCTGTTGCTGTGCTGATGCTGGAATAATGATGATGATGATGAT 1262
QY 1571 AAGCATTTTTTTTCTTTGAACCTTAAGGTTCTATATTAATAAGCAGACAGATGATGAT 1630
DB 1263 AAGCATTTTTTTTCTTTGAACCTTAAGGTTCTATATTAATAAGCAGACAGATGATGAT 1322
QY 1631 CATTTTATACATGAGAGATCTTCTTTGCTGATACACAGATGATGATGATGATGATGATGAT 1690
DB 1323 CATTTTATACATGAGAGATCTTCTTTGCTGATACACAGATGATGATGATGATGATGATGAT 1381
QY 1691 AAGAACTTTATGCTCCCTACGCTGCTGCTAAATATATCTAATTTCCAGATGCTTTGTAGA 1750
DB 1382 AAGAACTTTATGCTCCCTACGCTGCTGCTAAATATATCTAATTTCCAGATGCTTTGTAGA 1441
QY 1751 TGACGTGAAGTATTTGTGAGCCACATATTTGGAGTTCTAGATTGAGTGAATGGCAGAGAA 1810
DB 1442 TGACGTGAAGTATTTGTGAGCCACATATTTGGAGTTCTAGATTGAGTGAATGGCAGAGAA 1501
QY 1811 GGGCCATCTCCATTTGATGATGATTAAGTGAACCAAACTAGTTCTCGGAATTTCTACAGAGAA 1870
DB 1502 GGGCCATCTCCATTTGATGATGATTAAGTGAACCAAACTAGTTCTCGGAATTTCTACAGAGAA 1561
QY 1871 GGAGGAATCAGACTGAGAGAGCTGTGACATAGAGCTGAAGACCAAAAGCTTTGAATTT 1930

```

```

DB 1562 GGAGGAATCAGACTGARGAAGCTGTGACATGAGCTTGAAGACCAAAAGCTTTGAATTT 1621
QY 1931 TCGGAGCTGCTCATGTGATGATTAATACAGCTGCTCTTTATGAGTCAATCA 1990
DB 1622 TCGGAGCTGCTCATGTGATGATTAATACAGCTGCTCTTTATGAGTCAATCA 1681
QY 1991 TATTTTATGAGTGAATTAATTAAGAAAAAATTTACAA 2028
DB 1682 TATTTTATGAGTGAATTAATTAAGAAAAAATTTACAA 1719

```

```

RESULT 10
US-60-453-050-274
Sequence 274, Application US/60453050

```

```

: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele

```

```

: APPLICANT: LUKE, May
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

```

```

: TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001457

```

```

: CURRENT APPLICATION NUMBER: US/60/453,050
: CURRENT FILING DATE: 2003-03-10

```

```

: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 274

```

```

: LENGTH: 1719
: TYPE: DNA

```

```

: ORGANISM: Homo sapiens
: US-60-453-050-274

```

```

Query Match          30.5%; Score 622; DB 11; Length 1719;
Best Local Similarity 99.4%; Pred. No. 1.4e-298;
Matches 992; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1031 GTGCAAGATTTCATCCAGGGCCACAAAGCCTACAGCCACTCCCAATTAAGTGCCTATCAAT 1090
DB 723 GTGCAAAAGTTTCATCCAGGGCCACAAAGCCTACAGCCACTCCCAATTAAGTGCCTATCAAT 782
QY 1091 GAAGCTGAGCAAGAGAGTATACCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1150
DB 783 GAAGCTGAGCAAGAGAGTATACCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 842
QY 1151 GATCGGAGTGGGACGCGACATTAATAATCCCAATCCCACTTAACCACTGAAGAAAGA 1210
DB 843 GATCGGAGTGGGACGCGACATTAATAATCCCAATCCCACTTAACCACTGAAGAAAGA 902
QY 1211 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCCTCCAGACTAT 1270
DB 903 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCCTCCAGACTAT 962
QY 1271 AACCAAGAACTTAAGGAGAGGATCCCAAGGCGAGATGATCAAGAGCTGAATATCCAGC 1330
DB 963 AACCAAGAACTTAAGGAGAGGATCCCAAGGCGAGATGATCAAGAGCTGAATATCCAGC 1022
QY 1331 GTTTAAGACATGTCACATGGGCTTTGGAAGAGTGTGGGATCCAGTTCCAGAGAGAG 1390
DB 1023 GTTTAAGACATGTCACATGGGCTTTGGAAGAGTGTGGGATCCAGTTCCAGAGAGAG 1082
QY 1391 GGGTATGTTGTCCTCCACATCGGCGCAAGAGTGTGCGAATTTCTGTGCATAGCAG 1450
DB 1083 GGGTATGTTGTCCTCCACATCGGCGCAAGAGTGTGCGAATTTCTGTGCATAGCAG 1142
QY 1451 AAAAGCTCCACCATTTTCTTTGATGCTTTTAAGTCTACAGTTCTCTATATAGAA 1510
DB 1143 AAAAGCTCCACCATTTTCTTTGATGCTTTTAAGTCTACAGTTCTCTATATAGAA 1202
QY 1511 ACAGCAGTCTGTGACGTCCTGTTGCTGTGCTGATGCTGGAATAATGATGATGATGATGAT 1570
DB 1203 ACAGCAGTCTGTGACGTCCTGTTGCTGTGCTGATGCTGGAATAATGATGATGATGATGAT 1262
QY 1571 AAGCATTTTTTTTCTTTGAACCTTAAGGTTCTATATTAATAAGCAGACAGATGATGAT 1630

```

Query Match	30.5%	Score 622	DB 11	Length 2119
Best Local Similarity	99.4%	Pred. No. 1.4e-298		
Matches 992	Conservative	0	Mismatches 5	Indels 1
QY 1031	GTGCAAAAGTTTCATCCAGGCGCACAAAGCCTACAGCAGCTCCATATAAGATGCCATACANT	1090		
Db 1123	GTGCAAAAGTTTCATCCAGGCGCACAAAGCCTACAGCAGCTCCATATAAGATGCCATACANT	1182		
QY 1091	GAAGCTGGAACAAGAGAAGTTATCACCCTGTGTGACTCTGTGATCGAATCATATTGGG	1150		
Db 1183	GAAGCTGGAACAAGAGAAGTTATCACCCTGTGTGACTCTGTGATCGAATCATATTGGG	1242		
QY 1151	GATCGCGAATGGGCGCGCCACATAAATTCCAAATCCACTTGAAACCACTGAAGAAAGA	1210		
Db 1243	GATCGCGAATGGGCGCGCCACATAAATTCCAAATCCACTTGAAACCACTGAAGAAAGA	1302		
QY 1211	AGAAAGTTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCCTCCAGACTAT	1270		
Db 1303	AGAAAGTTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCCTCCAGACTAT	1362		
QY 1271	AACAAAGAAGCTAAAGGGAAGGATGCCAGGCGCGAATGATCAAGAGCTGAATATGCACC	1330		
Db 1363	AACAAAGAAGCTAAAGGGAAGGATGCCAGGCGCGAATGATCAAGAGCTGAATATGCACC	1422		

```

RESULT 12
US-60-453-050-275
Sequence 275, Application US/60453050
GENERAL INFORMATION:
APPLICANT: CARCILL, Michele
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001457
CURRENT APPLICATION NUMBER: US/60/453,050
CURRENT FILING DATE: 2003-03-10
NUMBER OF SEQ. ID NOS: 82762
SOFTWARE: FastSeq for Windows Version 4.0
SEQ. ID NO 275
LENGTH: 2119
TYPE: DNA
ORGANISM: Homo sapiens
US-60-453-050-275

```

Query Match	30.5%;	Score 622;	DB 11;	Length 2119;
Best Local Similarity	99.4%;	Pred. No. 1.4e-298;		
Matches 992;	Conservative 0;	Mismatches 5;	Indels 1;	Gaps 1;

```

OY 1031 GTGCAAGTTTCATCCAGGGCCACAAAGCTTACAGCCACTCCAAATAAGATGCCATACAT 1090
    |||||||
DB 1123 GTGCAAGTTTCATCCAGGGCCACAAAGCTTACAGCCACTCCAAATAAGATGCCATACAT 1182
OY 1091 GAACTGGAACACAGAAAGTTATACCTGTGTGACCTCTGTGATGCAATCATCTATGGG 1150
    |||||||
DB 1183 GAACTGGAACACAGAAAGTTATACCTGTGTGACCTCTGTGATGCAATCATCTATGGG 1242
OY 1151 GATCCGGAATGGGCGAGCCACATAAATCCAAATCCCACTTGAACCACTGAAGAAAGA 1210
    |||||||
DB 1243 GATCCGGAATGGGCGAGCCACATAAATCCAAATCCCACTTGAACCACTGAAGAAAGA 1302
OY 1211 AGAAGATTGAGCTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCACCACTAT 1270
    |||||||
DB 1303 AGAAGATTGAGCTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCACCACTAT 1362
OY 1271 AACAAAGAACCTAAAGGAAGGATCCCGAGGCGAATGATCAAGAGCTGAATGCAC 1330
    |||||||
DB 1363 AACAAAGAACCTAAAGGAAGGATCCCGAGGCGAATGATCAAGAGCTGAATGCAC 1422
OY 1331 GTTTAAGAGACATGTCCAGTGGCCCTTGGAAAGGTGTGGGATCCAGTTCAGAGAGAG 1390
    |||||||
DB 1423 GTTTAAGAGACATGTCCAGTGGCCCTTGGAAAGGTGTGGGATCCAGTTCAGAGAGAG 1482
OY 1391 GGGTATGTTTGTCTCCAGTGGGCAAGAGATGCTATGCGAATCTCTCATAGACG 1450
    |||||||
DB 1483 GGGTATGTTTGTCTCCAGTGGGCAAGAGATGCTATGCGAATCTCTCATAGACG 1542
OY 1451 AAAAGCTCCACACATTTCTTTTGTATGTGTTTAAAGTCTGACGTTCTATATAGAA 1510
    |||||||
DB 1543 AAAAGCTCCACACATTTCTTTTGTATGTGTTTAAAGTCTGACGTTCTATATAGAA 1602
OY 1511 ACAGAGGCTTGTGAGCTCCCTGTGTGTGCTGATGTCTGTGAATGAGATGTTCCAGA 1570
    |||||||
DB 1603 ACAGAGGCTTGTGAGCTCCCTGTGTGTGCTGATGTCTGTGAATGAGATGTTCCAGA 1662
OY 1571 AAGCATTTTTTTTTCTTTTGAACCTTAAAGTTCTATATTAATAAGCAGACAGATTCGA 1630
    |||||||
DB 1663 AAGCATTTTTTTTTCTTTTGAACCTTAAAGTTCTATATTAATAAGCAGACAGATTCGA 1722
OY 1631 CATTTTATATACATGAGATCTTCTTGTGTGTAATACAGAGATTGATGCTCATCCCTTAA 1690
    |||||||
DB 1723 CATTTTATATACATGAGATCTTCTTGTGTGTAATACAGAGATTGATGCTCATCCCTTAA 1781
OY 1691 AAGAGTTTATGTCCCTGACCTGTGCTAAATTTATCTAATTTCCAGATGCTTTGTAGA 1750
    |||||||
DB 1782 AAGAGTTTATGTCCCTGACCTGTGCTAAATTTATCTAATTTCCAGATGCTTTGTAGA 1841
OY 1751 TGACTGAAGTATTTGTGAGCCACATATTGGAGATTCTAGATTGAGTGAATGGCAGAAA 1810
    |||||||
DB 1842 TGACTGAAGTATTTGTGAGCCACATATTGGAGATTCTAGATTGAGTGAATGGCAGAAA 1901
OY 1811 GGGCCTCATCTCATAGATGATTAAGTGAACCAAACTACTCTCGGAATTTCTACAGAGAA 1870
    |||||||
DB 1902 GGGCCTCATCTCATAGATGATTAAGTGAACCAAACTACTCTCGGAATTTCTACAGAGAA 1961
OY 1871 GGAGGAATCAGATGAGGAAGCTGTGACATAGGACTTGAACCAAACTACTTTGAAT 1930
    |||||||
DB 1962 GGAGGAATCAGATGAGGAAGCTGTGACATAGGACTTGAACCAAACTACTTTGAAT 2021
OY 1931 TCGAGCTGCTCATGTGTGAGTATTAATCACTGCTGTCTTTCTATGAGTACAAATCTA 1990
    |||||||
DB 2022 TCGAGCTGCTCATGTGTGAGTATTAATCACTGCTGTCTTTCTATGAGTACAAATCTA 2081
OY 1991 TATTTTATGAAGTTTAAATTAAGAAAAAATTTACAA 2028
    |||||||
DB 2082 TATTTTATGAAGTTTAAATTAAGAAAAAATTTACAA 2119

```

RESULT 13  
 US-09-532-315B-3472  
 ; Sequence 3472, Application US/09532315B  
 ; GENERAL INFORMATION:

```

: APPLICANT: Sellhauer, Jeffrey J.
: APPLICANT: Delegeane, Angelo M.
: APPLICANT: Stuart, Susan G.
: APPLICANT: Stuve, Laura L.
: APPLICANT: Mullahy, Sara J.
: APPLICANT: Naughton, Rebecca E.
: TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
: FILE REFERENCE: PD-1002 CIP
: CURRENT APPLICATION NUMBER: US/09/532,315B
: CURRENT FILING DATE: 2000-03-24
: PRIOR APPLICATION NUMBER: 07/916,491
: PRIOR FILING DATE: 1992-07-17
: PRIOR APPLICATION NUMBER: 07/977,780
: PRIOR FILING DATE: 1992-11-19
: PRIOR APPLICATION NUMBER: 08/100,523
: PRIOR FILING DATE: 1993-08-03
: PRIOR APPLICATION NUMBER: 09/008,119
: PRIOR FILING DATE: 1998-01-16
: PRIOR APPLICATION NUMBER: 08/196,364
: PRIOR FILING DATE: 1994-02-14
: PRIOR APPLICATION NUMBER: 08/282,991
: PRIOR FILING DATE: 1994-07-28
: PRIOR APPLICATION NUMBER: 08/438,571
: PRIOR FILING DATE: 1995-05-10
: PRIOR APPLICATION NUMBER: 08/179,873
: PRIOR FILING DATE: 1994-01-11
: PRIOR APPLICATION NUMBER: 08/504,732
: PRIOR FILING DATE: 1995-07-20
: PRIOR APPLICATION NUMBER: 08/137,951
: PRIOR FILING DATE: 1993-10-14
: Remaining Prior Application data removed - See file wrapper or PALM.
: NUMBER OF SEQ ID NOS: 42212
: SOFTWARE: PERL Program
: SEQ ID NO: 3472
: LENGTH: 555
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Incyte ID No: hu01286535
US-09-532-315B-3472

Query Match          19.8%; Score 404; DB 6; Length 555;
Best Local Similarity 100.0%; Pred. No. 4,4e-190;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 14  
US-09-532-315B-3471  
Sequence 3471, Application US/09532315B  
GENERAL INFORMATION:  
APPLICANT: Sellhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES  
FILE REFERENCE: PD-1002 CIP  
CURRENT FILING DATE: 2000-03-24  
CURRENT APPLICATION NUMBER: US/09/532.315B  
PRIOR APPLICATION NUMBER: 07/916,491  
PRIOR FILING DATE: 1992-07-17  
PRIOR APPLICATION NUMBER: 07/977,780  
PRIOR FILING DATE: 1992-11-19  
PRIOR APPLICATION NUMBER: 08/100,523  
PRIOR FILING DATE: 1993-08-03  
PRIOR APPLICATION NUMBER: 09/008,119  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 08/196,364  
PRIOR FILING DATE: 1994-02-14  
PRIOR APPLICATION NUMBER: 08/282,991  
PRIOR FILING DATE: 1994-07-28  
PRIOR APPLICATION NUMBER: 08/438,571  
PRIOR FILING DATE: 1995-05-10  
PRIOR APPLICATION NUMBER: 08/179,873  
PRIOR FILING DATE: 1994-01-11  
PRIOR APPLICATION NUMBER: 08/504,732  
PRIOR FILING DATE: 1995-07-20  
PRIOR APPLICATION NUMBER: 08/137,951  
PRIOR FILING DATE: 1993-10-14  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 42212  
SOFTWARE: PERL Program  
SEQ ID NO 3471  
LENGTH: 392  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: hu01253362  
US-09-532-315B-3471

Query Match 13.2%; Score 269; DB 6; Length 392;  
Best Local Similarity 99.7%; Pred. No. 6,7e-123;  
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1690 AAGAACTTTTATGTCCTGACCTGCTGCTAAATTTATCTAATTTCCAGATGCTTTGTAG 1749  
DB 69 AAGAACTTTTATGTCCTGACCTGCTGCTAAATTTATCTAATTTCCAGATGCTTTGTAG 128  
QY 1750 ATGACTGAATTTTGTGACCCACATATTTGGAGTTCTAATTTGATGATGCGAGAA 1809  
DB 129 ATGACTGAATTTTGTGACCCACATATTTGGAGTTCTAATTTGATGATGCGAGAA 188  
QY 1810 AGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCGGAATTTCTACAGAGA 1869  
DB 189 AGGCGCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCGGAATTTCTACAGAGA 248  
QY 1870 AGAGGGAATCAGACTGAGAGAGCTGTGACATAGGAGTTGAAGCAACCAAGCTTTGAAT 1929  
DB 249 AGAGGGAATCAGACTGAGAGAGCTGTGACATAGGAGTTGAAGCAACCAAGCTTTGAAT 308  
QY 1930 TTGGAGCTGCTCATGCTGATGATTAATATACATGCTCTTTCTTATGATGATTAACATCT 1989  
DB 309 TTGGAGCTGCTCATGCTGATGATTAATATACATGCTCTTTCTTATGATGATTAACATCT 368  
QY 1990 ATATTTTATGAAGTTAA 2009

DB 369 ATATTTTATGAAGTTAA 388  
RESULT 15  
US-09-532-315B-9521  
Sequence 9521, Application US/09532315B  
GENERAL INFORMATION:  
APPLICANT: Sellhamer, Jeffrey J.  
APPLICANT: Delegeane, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES  
FILE REFERENCE: PD-1002 CIP  
CURRENT FILING DATE: 2000-03-24  
CURRENT APPLICATION NUMBER: 07/916,491  
PRIOR APPLICATION NUMBER: 07/977,780  
PRIOR FILING DATE: 1992-11-19  
PRIOR APPLICATION NUMBER: 08/100,523  
PRIOR FILING DATE: 1993-08-03  
PRIOR APPLICATION NUMBER: 09/008,119  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 08/196,364  
PRIOR FILING DATE: 1994-02-14  
PRIOR APPLICATION NUMBER: 08/282,991  
PRIOR FILING DATE: 1994-07-28  
PRIOR APPLICATION NUMBER: 08/438,571  
PRIOR FILING DATE: 1995-05-10  
PRIOR APPLICATION NUMBER: 08/179,873  
PRIOR FILING DATE: 1994-01-11  
PRIOR APPLICATION NUMBER: 08/504,732  
PRIOR FILING DATE: 1995-07-20  
PRIOR APPLICATION NUMBER: 08/137,951  
PRIOR FILING DATE: 1993-10-14  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 42212  
SOFTWARE: PERL Program  
SEQ ID NO 9521  
LENGTH: 293  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte ID No: hu00875621  
US-09-532-315B-9521

Query Match 13.0%; Score 265; DB 6; Length 293;  
Best Local Similarity 100.0%; Pred. No. 6,6e-121;  
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GTGGGAGGTGGGCTCAGGGGCTGCAACGAGACCTACCTCTTGTAGATTTCCGGGGCC 103  
DB 29 GTGGGAGGTGGGCTCAGGGGCTGCAACGAGACCTACCTCTTGTAGATTTCCGGGGCC 88  
QY 104 ACGGGACCGCGCAATTCACAGCTGGGGTTGACAGTGGCCAGCGGCTCGGGCTGAGATC 163  
DB 89 ACGGGACCGCGCAATTCACAGCTGGGGTTGACAGTGGCCAGCGGCTCGGGCTGAGATC 148  
QY 164 GTGAGGCTGATCTCCATGACAGGCTCTATGAAGGCTTGAACATCTACCAACCAAGCTTTCT 223  
DB 149 GTGAGGCTGATCTCCATGACAGGCTCTATGAAGGCTTGAACATCTACCAACCAAGCTTTCT 208  
QY 224 GCCCAGAGAGAGAAATTCGCGGACCAACATGATGAGTTGTGATCCTCTTGAGAC 283  
DB 209 GCCCAGAGAGAGAAATTCGCGGACCAACATGATGAGTTGTGATCCTCTTGAGAC 268  
QY 284 AATTACACAGTGTGAGCTTCAGAA 308  
DB 269 AATTACACAGTGTGAGCTTCAGAA 293

Tue Apr 22 09:11:14 2003

us-09-513-151-3.oli.rnpn

Page 11

Search completed: April 22, 2003, 00:01:33  
Job time : 977 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:32:57 ; Search time 5186 Seconds

(Without alignments)  
9895.067 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041

Sequence: 1 CTCGCATTAACATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 24791104 seqs, 12571243825 residues

Word size : 0

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

Pending_Patents_MA_Main:
1: /cgn2_6/ptodata/1/pna/US0950_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

```

Result No.	Score	Query Match	Length	DB ID	Description
1	2041	100.0	2041	19 US-09-513-151-3	Sequence 3, Appl
2	939	46.0	2108	40 US-10-113-013-165	Sequence 155, App
3	939	46.0	2129	1 PCT-US02-07826-152	Sequence 152, App
4	939	46.0	2129	39 US-10-097-340-152	Sequence 152, App
5	918	45.0	2109	76 US-60-324-185-33690	Sequence 33690, A
6	823	40.3	2129	1 PCT-US02-07826-150	Sequence 150, App
7	823	40.3	2129	39 US-10-097-340-150	Sequence 150, App
8	818	40.1	2101	61 US-60-112-360-26034	Sequence 26034, A
9	778	38.1	1183	29 US-09-757-028-874	Sequence 874, App
10	778	38.1	1183	42 US-10-222-911-874	Sequence 874, App
11	757	37.1	1041	16 US-09-205-070-8140	Sequence 8140, Ap
12	757	37.1	1041	17 US-09-340-623-8140	Sequence 8140, Ap
13	757	37.1	1041	33 US-09-888-888-8140	Sequence 8140, Ap
14	757	37.1	1041	33 US-09-888-888-8140	Sequence 8140, Ap
15	751	36.8	1749	1 PCT-US02-18947-1692	Sequence 1692, Ap
16	751	36.8	1749	41 US-10-172-118-1692	Sequence 1692, Ap
17	697	34.1	1223	28 US-09-705-256A-6006	Sequence 6006, Ap
18	697	34.1	1223	60 US-60-164-285-6006	Sequence 6006, Ap
19	696	34.1	1039	17 US-09-359-822-3465	Sequence 3465, Ap
20	696	34.1	1039	17 US-09-359-822-3465	Sequence 3465, Ap
21	696	34.1	1039	34 US-09-919-002-3465	Sequence 3465, Ap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

22 694 34.0 1858 18 US-09-471-275-2814  
 23 638 31.3 1009 17 US-09-399-932-5314  
 24 480 23.5 519 25 US-09-652-121-891  
 25 473 23.2 717 1 PCT-US01-01239-494  
 26 473 23.2 717 1 PCT-US01-01349-166  
 27 473 23.2 717 1 PCT-US01-01349-364  
 28 473 23.2 717 30 US-09-764-853-166  
 29 473 23.2 717 30 US-09-764-853-364  
 30 473 23.2 717 30 US-09-764-902-494  
 31 473 23.2 717 30 US-10-072-326-166  
 32 473 23.2 717 30 US-10-072-326-364  
 33 459 22.5 508 25 US-09-652-127-7656  
 34 457 22.4 1069 25 US-09-652-124-7628  
 35 447 21.9 487 16 US-09-235-076-20244  
 36 447 21.9 487 16 US-09-289-768-25904  
 37 447 21.9 487 17 US-09-332-782-20244  
 38 447 21.9 487 29 US-09-737-223-20244  
 39 447 21.9 487 34 US-09-918-995-20244  
 40 447 21.9 487 35 US-09-939-397-25904  
 41 443 21.7 636 22 US-09-584-852-5059  
 42 438 21.5 480 16 US-09-289-768-25903  
 43 438 21.5 480 35 US-09-939-397-25903  
 44 431 21.1 487 18 US-09-489-036-15715  
 45 431 21.1 487 35 US-09-943-143-15715

## ALIGNMENTS

Sequence 2814, Ap  
 Sequence 5314, Ap  
 Sequence 891, App  
 Sequence 494, App  
 Sequence 166, App  
 Sequence 364, App  
 Sequence 166, App  
 Sequence 494, App  
 Sequence 166, App  
 Sequence 364, App  
 Sequence 7656, App  
 Sequence 7628, App  
 Sequence 20244, A  
 Sequence 25904, A  
 Sequence 20244, A  
 Sequence 20244, A  
 Sequence 20244, A  
 Sequence 5059, App  
 Sequence 25903, A  
 Sequence 25903, A  
 Sequence 15715, A  
 Sequence 15715, A

RESULT 1  
 US-09-513-151-3  
 Sequence 3, Application US/09513151  
 GENERAL INFORMATION:  
 APPLICANT: MCGILL UNIVERSITY  
 TITLE OF INVENTION: THE C. ELEGANS gro-1 GENE  
 NUMBER OF SEQUENCES: 62  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: SMARBY OGILVY RENAULT  
 STREET: 1981 McGill College Avenue - Suite 1600  
 CITY: Montreal  
 STATE: QC  
 COUNTRY: Canada  
 ZIP: H3A 2Y3  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows  
 SOFTWARE: FASTSEQ for Windows Version 2.0b  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/513,151  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/CA98/00803  
 FILING DATE: 20-AUG-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: CA 2,210,251  
 FILING DATE: 25-AUG-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ctl, France  
 REGISTRATION NUMBER: 4166  
 REFERENCE/DOCKET NUMBER: 1770-179\*US\* FC/9C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 514 845-7126  
 TELEFAX: 514 288-8319  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2041 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: Genomic DNA

US-09-513-151-3  
 Query Match 100.0%; Score 2041; DB 19; Length 2041;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCCATTAAGATGGCGCGCTGGCGCTGCACAGAGAGTTCCTGGGCGAGTGGCTCAG 60  
 DB 1 CTGCCATTAAGATGGCGCGCTGGCGCTGCACAGAGAGTTCCTGGGCGAGTGGCTCAG 60  
 QY 61 GGGCTGCAACGAGACCTTACCTTGTAGTATTCGCGGGCCAGCGGACCGCAAAATC 120  
 DB 61 GGGCTGCAACGAGACCTTACCTTGTAGTATTCGCGGGCCAGCGGACCGCAAAATC 120  
 QY 121 CACGCTGGCGCTTCAGCTAGGCGGCTGGCGGCTGAGATGTCAGGCGTACTCAT 180  
 DB 121 CACGCTGGCGCTTCAGCTAGGCGGCTGGCGGCTGAGATGTCAGGCGTACTCAT 180  
 QY 181 GCAAGTCTATGAGAGGCTTACATCATCACCACCAAGGTTTCGCCCAAGCAGAGAAAT 240  
 DB 181 GCAAGTCTATGAGAGGCTTACATCATCACCACCAAGGTTTCGCCCAAGCAGAGAAAT 240  
 QY 241 CTGCCGACACCATGATGATGCTTGTGATCCTTGTGACCAATTAACAGTGGTGA 300  
 DB 241 CTGCCGACACCATGATGATGCTTGTGATCCTTGTGACCAATTAACAGTGGTGA 300  
 QY 301 CTTCAGAAATAGACCACTGCTCTGATTAAGATATTTGCCGAGACAAATTCCTAT 360  
 DB 301 CTTCAGAAATAGACCACTGCTCTGATTAAGATATTTGCCGAGACAAATTCCTAT 360  
 QY 361 TGTGTTGGGAGGAGCAATTTATTAATGATCTTGTCTGTGAAAGTTCTTCAATAC 420  
 DB 361 TGTGTTGGGAGGAGCAATTTATTAATGATCTTGTCTGTGAAAGTTCTTCAATAC 420  
 QY 421 CAAGCCCCAGAGATGAGGAGCTGAGAAAGATTTGACCCGAAAGTGGAGTGAAGAAG 480  
 DB 421 CAAGCCCCAGAGATGAGGAGCTGAGAAAGATTTGACCCGAAAGTGGAGTGAAGAAG 480  
 QY 481 GGATGCTCTTGTACTTACAAACGCTTGAAGCGTGGAGCCGAAAGTGGTCCAGCT 540  
 DB 481 GGATGCTCTTGTACTTACAAACGCTTGAAGCGTGGAGCCGAAAGTGGTCCAGCT 540  
 QY 541 GCATCCACATGAGAAACGCAAAATGCGCCAGGACTTGCAGTTTGAAGAAACAGAAAT 600  
 DB 541 GCATCCACATGAGAAACGCAAAATGCGCCAGGACTTGCAGTTTGAAGAAACAGAAAT 600  
 QY 601 CTCTCATAGTGAATTTCTCCATGCTCAACATACAGAAAGGAGTGGTCCCTTGGAG 660  
 DB 601 CTCTCATAGTGAATTTCTCCATGCTCAACATACAGAAAGGAGTGGTCCCTTGGAG 660  
 QY 661 TCCTCTGAAGTCTTCAACCTTGATCTTTGGCTTCATGCTGACAGGAGGATTTAGA 720  
 DB 661 TCCTCTGAAGTCTTCAACCTTGATCTTTGGCTTCATGCTGACAGGAGGATTTAGA 720  
 QY 721 TGAGCCCTTGGATTAAGAGGCTGATGATGATGCTGCTGGGCTTGGAGAACTAG 780  
 DB 721 TGAGCCCTTGGATTAAGAGGCTGATGATGATGATGCTGCTGGGCTTGGAGAACTAG 780  
 QY 781 AGATTTTACAGAGGCTTATATATGAGAAATGTTGGAAATATGACAGGATCTACAA 840  
 DB 781 AGATTTTACAGAGGCTTATATATGAGAAATGTTGGAAATATGACAGGATCTACAA 840  
 QY 841 TGGTATCTTCAATCAATGATGCTTCAAGAAATTTCAAGATACCTGATCAGGAGGANA 900  
 DB 841 TGGTATCTTCAATCAATGATGCTTCAAGAAATTTCAAGATACCTGATCAGGAGGANA 900  
 QY 901 ATGCAACACTGAGACCTAGTAAAGGAGTCTTAAAGAAAGACCTGATGCTGATGCTG 960  
 DB 901 ATGCAACACTGAGACCTAGTAAAGGAGTCTTAAAGAAAGACCTGATGCTGATGCTG 960  
 QY 961 TGTCTATGCTTAAAGAGTATGATGCTGCAAGTGGAGAGTCTGTTTGAACCTGC 1020  
 DB 961 TGTCTATGCTTAAAGAGTATGATGCTGCAAGTGGAGAGTCTGTTTGAACCTGC 1020



Oy 1021 TCTTGAATGCTGCAAAAGTTTCATCCAGGGCCACAGCCTACAGCCTCCCAATTAAGAT 1080  
 Db 1021 TCTTGAATGCTGCAAAAGTTTCATCCAGGGCCACAGCCTACAGCCTCCCAATTAAGAT 1080  
 Oy 1081 GCCATTAAGTAAAGCTGAGAAACAAGAAAGTATACCTCTGTGACCTCTGTGATCGAAT 1140  
 Db 1081 GCCATTAAGTAAAGCTGAGAAACAAGAAAGTATACCTCTGTGACCTCTGTGATCGAAT 1140  
 Oy 1141 CATCATTTGGGATCGCGAATGGGACCCACATATAATCCAAATCCCACTTGAAACCACT 1200  
 Db 1141 CATCATTTGGGATCGCGAATGGGACCCACATATAATCCAAATCCCACTTGAAACCACT 1200  
 Oy 1201 GAAAGAAAAGAAAGATTTGGACTGAGATGCTGTCAACCCATAGAAAGTCAAGAGTTC 1260  
 Db 1201 GAAAGAAAAGAAAGATTTGGACTGAGATGCTGTCAACCCATAGAAAGTCAAGAGTTC 1260  
 Oy 1261 CCCAGACTATTAACAAGAAAGCTTAAAGGAGGATCCCAAGGAGCAAGATGATCAAGACT 1320  
 Db 1261 CCCAGACTATTAACAAGAAAGCTTAAAGGAGGATCCCAAGGAGCAAGATGATCAAGACT 1320  
 Oy 1321 GAAATGACGCTTTAAGAGACATGTCAGTGCCCTTTGGAAGGTGTGGGATCCAGTT 1380  
 Db 1321 GAAATGACGCTTTAAGAGACATGTCAGTGCCCTTTGGAAGGTGTGGGATCCAGTT 1380  
 Oy 1381 CAGAGGAGGAGGATGTTGTTGCTCCAGTCTGGGCAAGAGATGCTATGCGAATTCCTC 1440  
 Db 1381 CAGAGGAGGAGGATGTTGTTGCTCCAGTCTGGGCAAGAGATGCTATGCGAATTCCTC 1440  
 Oy 1441 TGCATAGCAGAAAAGCTCCCAACATTTCTTTGATGTGTTTAAAGTCTCACCTTC 1500  
 Db 1441 TGCATAGCAGAAAAGCTCCCAACATTTCTTTGATGTGTTTAAAGTCTCACCTTC 1500  
 Oy 1501 TATAATAGAAACAGCAGCTCTTGTGACGCTCTGTGTGCGTGAATGTCTGGAATGATG 1560  
 Db 1501 TATAATAGAAACAGCAGCTCTTGTGACGCTCTGTGTGCGTGAATGTCTGGAATGATG 1560  
 Oy 1561 TAGTTCAGGAAAGCAATTTTCTTTGTAACCTTAAAGGTCTATTAATTAAGAGCAGC 1620  
 Db 1561 TAGTTCAGGAAAGCAATTTTCTTTGTAACCTTAAAGGTCTATTAATTAAGAGCAGC 1620  
 Oy 1621 ACAGATTCACATTTTATACATGAGATCTTTGCTGTGCTGAATACAGATGATGCTC 1680  
 Db 1621 ACAGATTCACATTTTATACATGAGATCTTTGCTGTGCTGAATACAGATGATGCTC 1680  
 Oy 1681 ATCCCTTTAAAGAAAGTATGTCCTGACCTGTGCTAAATTAATTAATTTCCAGATG 1740  
 Db 1681 ATCCCTTTAAAGAAAGTATGTCCTGACCTGTGCTAAATTAATTAATTTCCAGATG 1740  
 Oy 1741 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 Db 1741 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800  
 Oy 1801 TGGCAGGAAAGGAGGATCCATCCATGATGATGATGATGATGATGATGATGATGATG 1860  
 Db 1801 TGGCAGGAAAGGAGGATCCATCCATGATGATGATGATGATGATGATGATGATGATG 1860  
 Oy 1861 CTACAGAGAGAGGAGGATCCATCCATGATGATGATGATGATGATGATGATGATGATG 1920  
 Db 1861 CTACAGAGAGAGGAGGATCCATCCATGATGATGATGATGATGATGATGATGATGATG 1920  
 Oy 1921 CTTTGAATTTTCCGAGCTGCTCATGTGTGATGATGATGATGATGATGATGATGATG 1980  
 Db 1921 CTTTGAATTTTCCGAGCTGCTCATGTGTGATGATGATGATGATGATGATGATGATG 1980  
 Oy 1981 TACAATCTATATTTTATGAAAGTTTAAATTAAGAAAAATTTTCAAGAAAAA 2040  
 Db 1981 TACAATCTATATTTTATGAAAGTTTAAATTAAGAAAAATTTTCAAGAAAAA 2040  
 Oy 2041 A 2041  
 Db 2041 A 2041

RESULT 2  
 US-10-133-013-165  
 : Sequence 165, Application US/10133013  
 : GENERAL INFORMATION:  
 : APPLICANT: Bandman, Anna  
 : APPLICANT: Astromoff, Anna  
 : APPLICANT: Bandman, Olga  
 : APPLICANT: Cocks, Benjamin G.  
 : TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE  
 : FILE REFERENCE: PA-0049 US  
 : CURRENT APPLICATION NUMBER: US/10/133,013  
 : CURRENT FILING DATE: 2002-04-25  
 : PRIOR APPLICATION NUMBER: 60/287,067  
 : PRIORITY FILING DATE: 2001-04-27  
 : NUMBER OF SEQ ID NOS: 271  
 : SOFTWARE: PERL Program  
 : SEQ ID NO 165  
 : LENGTH: 2108  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: misc-feature  
 : OTHER INFORMATION: Incyte ID No: 977391.15  
 : US-10-133-013-165  
 Query Match 46.0%; Score 939; DB 40; Length 2108;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 CTGCCATTAAGATGCGCTCCGTCGCGGCTGACAGACAGTCTCTGTGGGAGTGGGCTCAG 60  
 Db 1 CTGCCATTAAGATGCGCTCCGTCGCGGCTGACAGACAGTCTCTGTGGGAGTGGGCTCAG 60  
 Oy 61 GGGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 Db 61 GGGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 Oy 121 CAGCTGCGGTTGACAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
 Db 121 CAGCTGCGGTTGACAGTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
 Oy 181 GCAGGTATGAGAGGCTTACAGATCATCACCAAGAGTTTGTGCCCAAGACAGAGAT 240  
 Db 181 GCAGGTATGAGAGGCTTACAGATCATCACCAAGAGTTTGTGCCCAAGACAGAGAT 240  
 Oy 241 CTGCCGAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 Db 241 CTGCCGAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
 Oy 301 CTTCAAGAAATAGAGCACTGCTGTGATGATGATGATGATGATGATGATGATGATGATG 360  
 Db 301 CTTCAAGAAATAGAGCACTGCTGTGATGATGATGATGATGATGATGATGATGATGATG 360  
 Oy 361 TGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
 Db 361 TGTGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420  
 Oy 421 CAAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 Db 421 CAAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 Oy 481 GATGAGTCTTGTACTTACAAAGGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 Db 481 GATGAGTCTTGTACTTACAAAGGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 Oy 541 GCATCCACATGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 Db 541 GCATCCACATGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 Oy 601 CTCTCATAGTGAATTTCTCATGCTGACATACGGAAGAGGAGGAGGAGGAGGAGGAGGAGG 660  
 Db 601 CTCTCATAGTGAATTTCTCATGCTGACATACGGAAGAGGAGGAGGAGGAGGAGGAGGAGG 660



```

: CURRENT FILING DATE: 2002-03-14
: PRIOR APPLICATION NUMBER: 60/276,025
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 60/325,149
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 60/276,026
: PRIOR FILING DATE: 2001-03-14
: PRIOR APPLICATION NUMBER: 60/324,967
: PRIOR FILING DATE: 2001/09/26
: PRIOR APPLICATION NUMBER: 60/311,732
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: 60/325,102
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 60/323,580
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ ID NOS: 363
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 152
: LENGTH: 2129
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-097-340-152

```

Query Match 46.0%; Score 939; DB 39; Length 2129;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 939; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 CTGCGATTAATGCGCGTCCGTCGCGGTGACGACGAGCTTCCTGTGGGCGAGTGGCGTCAG 60
DB 5 CTCGCATTAAGATGCGCGTCCGTCGCGGTGACGACGAGCTTCCTGTGGGCGAGTGGCGTCAG 64
OY 61 GGGCGTGCACAGGACCTTACTCTCTGTAGTATCTGGGGCCACGGGCAACGCAATC 120
DB 65 GGGCGTGCACAGGACCTTACTCTCTGTAGTATCTGGGGCCACGGGCAACGCAATC 124
OY 121 CACGCTGCGCTTGACGCTAGGCGCAGCGCTCGCGGTGAGATGCTACGCGCTGACTCAT 180
DB 125 CACGCTGCGCTTGACGCTAGGCGCAGCGCTCGCGGTGAGATGCTACGCGCTGACTCAT 184
OY 181 GCAGGTCTATGAGAGCGCTAGACATCATCACCAAGCTTTCTGCCCAAGCAGAGAAAT 240
DB 185 GCAGGTCTATGAGAGCGCTAGACATCATCACCAAGCTTTCTGCCCAAGCAGAGAAAT 244
OY 241 CTGCGGCGACCAACATGATGATGCTTTGGATCTCTGTGCAATATACAGAGTGGGA 300
DB 245 CTGCGGCGACCAACATGATGATGCTTTGGATCTCTGTGCAATATACAGAGTGGGA 304
OY 301 CTTCAGAAATAGACCACTGCTCTGATTGAAGATATATTTGCCGAGACAAATTCCTAT 360
DB 305 CTTCAGAAATAGACCACTGCTCTGATTGAAGATATATTTGCCGAGACAAATTCCTAT 364
OY 361 TGTGTGGGAGAACCAATTATTAATTAATCTCTGCTGGAAGTTCTTGTCAATAC 420
DB 365 TGTGTGGGAGAACCAATTATTAATTAATCTCTGCTGGAAGTTCTTGTCAATAC 424
OY 421 CAACCCCGAGAGATGGGACTGGAAGATGATGACGGAATGAGGCTTGAAGGA 480
DB 425 CAACCCCGAGAGATGGGACTGGAAGATGATGACGGAATGAGGCTTGAAGGA 484
OY 481 GGAAGTCTTGTACTTCAACCAACGCTTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCT 540
DB 485 GGAAGTCTTGTACTTCAACCAACGCTTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCT 544
OY 541 GCATCCACATGACAAAGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGAT 600
DB 545 GCATCCACATGACAAAGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGAT 604
OY 601 CTCCTCATAGTAATTTCTCCATCGTCAACATAGGGAAGAGTGGTGGCTTCCCTTGAGG 660
DB 605 CTCCTCATAGTAATTTCTCCATCGTCAACATAGGGAAGAGTGGTGGCTTCCCTTGAGG 664
OY 661 TCCCTCATAGTAATTTCTCCATCGTCAACATAGGGAAGAGTGGTGGCTTCCCTTGAGG 720
DB 665 TCCCTCATAGTAATTTCTCCATCGTCAACATAGGGAAGAGTGGTGGCTTCCCTTGAGG 724

```

```

DB 665 TCCCTCATAGTAATTTCTCCATCGTCAACATAGGGAAGAGTGGTGGCTTCCCTTGAGG 724
OY 721 TGAGCGCTTGATTAAGAGGCTGATGATGCTTGTGCTGGGCTTGTGGAGAACTAAG 780
DB 725 TGAGCGCTTGATTAAGAGGCTGATGATGCTTGTGCTGGGCTTGTGGAGAACTAAG 784
OY 781 AGATTTTCAACAGCGCTTAATATCAGAGAAGATGTTGGGAAATATAGCCAGACTATCAAC 840
DB 785 AGATTTTCAACAGCGCTTAATATCAGAGAAGATGTTGGGAAATATAGCCAGACTATCAAC 844
OY 841 TGGTATCTTCCATCAATTAAGGCTTCAAGAAATTTACAGAGTACCTGATCACTGAGGAA 900
DB 845 TGGTATCTTCCATCAATTAAGGCTTCAAGAAATTTACAGAGTACCTGATCACTGAGGAA 904
OY 901 ATGCACACTGAGACTAGTAACCAAGCTTCTTAAGAAAG 939
DB 905 ATGCACACTGAGACTAGTAACCAAGCTTCTTAAGAAAG 943

```

RESULT 5

US-60-324-185-33690

Sequence 33690, Application us/60324185

GENERAL INFORMATION:

APPLICANT: Morris, MacDonald

APPLICANT: Lai, Preethi

TITLE OF INVENTION: METHOD FOR THE IDENTIFICATION OF SEQUENCE POLYMORPHISMS US

TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCE DATABASES, AND SINGLE NUCLEOTIDE

TITLE OF INVENTION: POLYMORPHISMS IDENTIFIED THEREBY

FILE REFERENCE: GX-0019-1 P

CURRENT APPLICATION NUMBER: US/60/324,185

CURRENT FILING DATE: 2001-09-21

NUMBER OF SEQ ID NOS: 35862

SOFTWARE: PERL Program

SEQ ID NO: 33690

LENGTH: 2109

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Inocyte ID No: 977391.2

NAME/KEY: unsure

LOCATION: 2106

OTHER INFORMATION: a, t, c, g, or other

US-60-324-185-33690

Query Match 45.0%; Score 918; DB 76; Length 2109;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 22 GCGCGCTGCAGAGCACTTCTGCGGAGCTGAGGCTCAGGGGCTGCAAGCAACCTACC 81
DB 20 GCGCGCTGCAGAGCACTTCTGCGGAGCTGAGGCTCAGGGGCTGCAAGCAACCTACC 79
OY 82 TCTTGTAGTATTCGCGGCGCACGCGGCGCAACGCTGCGTTCAGCTAGG 141
DB 80 TCTTGTAGTATTCGCGGCGCACGCGGCGCAACGCTGCGTTCAGCTAGG 139
OY 142 CCAAGCGCTGCGGCTGAGATCGTCAGCGCTGATCCATGACAGTCTATGAAGGCTAGA 201
DB 140 CCAAGCGCTGCGGCTGAGATCGTCAGCGCTGATCCATGACAGTCTATGAAGGCTAGA 199
OY 202 CATCATCACAACAAGGTTTCTGCCCAAGAGCAGAGAATGCTGCCGACCAATGATCAG 261
DB 200 CATCATCACAACAAGGTTTCTGCCCAAGAGCAGAGAATGCTGCCGACCAATGATCAG 259
OY 262 CTTTGTGATCTCTTGTGACCAATTAACAGTGGTGGACTTCAAGAAATAGCAACTGC 321
DB 260 CTTTGTGATCTCTTGTGACCAATTAACAGTGGTGGACTTCAAGAAATAGCAACTGC 319
OY 322 TCTGATTAAGATATATTTGGCCGAGCAAAATTTCTATTGTTGTGGAGGAACCAATTA 381
DB 320 TCTGATTAAGATATATTTGGCCGAGCAAAATTTCTATTGTTGTGGAGGAACCAATTA 379

```

```

Qy 382 TTACATGATCTCTCTCTGGAAGTTCTTGCATACCAAGCCAGAGATGGGCAC 441
    |||
Db 380 TTACATGATCTCTCTCTGGAAGTTCTTGCATACCAAGCCAGAGATGGGCAC 439
Qy 442 TGAGAAAGTATGACCGAAAGTGGAGCTTGAAGAGAGATGGTCTTGTACTTACAA 501
    |||
Db 440 TGAGAAAGTATGACCGAAAGTGGAGCTTGAAGAGAGATGGTCTTGTACTTACAA 499
Qy 502 AGCCTTAAGCCAGGTGGAGCCCGAAAGTGGCTCCCAAGCTGCATCCACATGACAAACGCAA 561
    |||
Db 500 AGCCTTAAGCCAGGTGGAGCCCGAAAGTGGCTCCCAAGCTGCATCCACATGACAAACGCAA 559
Qy 562 AGTGCCAGAGCTTGCAGATTTTGAAGAAACAGAGATCTCTCATAGTGAATTTCTCCA 621
    |||
Db 560 AGTGCCAGAGCTTGCAGATTTTGAAGAAACAGAGATCTCTCATAGTGAATTTCTCCA 619
Qy 622 TCGTAAACATACGGAAGAGTGTGTCCCTTGGAGGTCTCTGAAGTTCTTAAACC 681
    |||
Db 620 TCGTAAACATACGGAAGAGTGTGTCCCTTGGAGGTCTCTGAAGTTCTTAAACC 679
Qy 682 TTGCATCTTGGCTTGCATGCGGACGACGAGCTTATAGATAGCCCTTGGATAGAGGT 741
    |||
Db 680 TTGCATCTTGGCTTGCATGCGGACGACGAGCTTATAGATAGCCCTTGGATAGAGGT 739
Qy 742 GGATGACATGCTTGTCTGCTGGGCTTGTGAGAGAACTAGAGATTTTACAGACGCTATTA 801
    |||
Db 740 GGATGACATGCTTGTCTGCTGGGCTTGTGAGAGAACTAGAGATTTTACAGACGCTATTA 799
Qy 802 TCAGAAAGATTTTGGAAATATAGCCAGACTATCAACATGATCTTCCATCAATTTGG 861
    |||
Db 800 TCAGAAAGATTTTGGAAATATAGCCAGACTATCAACATGATCTTCCATCAATTTGG 859
Qy 862 CTTCAAGAAATTTTACAGATCTCTGATCAGTACGAGGAAATGACACTGGAGACTAGTAA 921
    |||
Db 860 CTTCAAGAAATTTTACAGATCTCTGATCAGTACGAGGAAATGACACTGGAGACTAGTAA 919
Qy 922 CCAGCTTCTAAAGAAAG 939
    |||
Db 920 CCAGCTTCTAAAGAAAG 937

```

```

RESULT 6
PCT-US02-07826-150
; Sequence 150, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 2129
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-07826-150

```

```

Query Match 40.3%; Score 823; DB 1; Length 2129;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 823; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 AATCCACGCTGGCGTTGACAGCTAGAGCCAGCGCTGGCGGTAGATCTGCACGCTGACT 176
    |||
Db 121 AATCCACGCTGGCGTTGACAGCTAGAGCCAGCGCTGGCGGTAGATCTGCACGCTGACT 180
Qy 177 CCATGACAGCTATATAGAGCCCTAGACATCATGACCAACAGGTTCTGCCCAAGCCAGA 236
    |||
Db 181 CCATGACAGCTATATAGAGCCCTAGACATCATGACCAACAGGTTCTGCCCAAGCCAGA 240
Qy 237 GAATTCGCCGCGCACCATATGATCAGCTTGTGTGATCTCTTGTGACCAATTCACAGTGG 296
    |||
Db 241 GAATTCGCCGCGCACCATATGATCAGCTTGTGTGATCTCTTGTGACCAATTCACAGTGG 300
Qy 297 TGGACTTGAAGAAATAGACCACTGCTGTGATTTGAAGATATATTTCCCGAGACAAATTC 356
    |||
Db 301 TGGACTTGAAGAAATAGACCACTGCTGTGATTTGAAGATATATTTCCCGAGACAAATTC 360
Qy 357 CTATTGTTGTGGAGAGCAACCAATTATATCATTTGAATCTCTGCTGGAAGTTCTTGCA 416
    |||
Db 361 CTATTGTTGTGGAGAGCAACCAATTATATCATTTGAATCTCTGCTGGAAGTTCTTGCA 420
Qy 417 ATACCAAGCCCGAGAGATGGGCACTGAGAAAGTGTATTTGACCGAAAGTGGAGCTTGAA 476
    |||
Db 421 ATACCAAGCCCGAGAGATGGGCACTGAGAAAGTGTATTTGACCGAAAGTGGAGCTTGAA 480
Qy 477 AGGAGATGATCTTGTACTTTCACAAAGCCCTTAAAGCAGGTGAGCCAGCAATTCGCTGCA 536
    |||
Db 481 AGGAGATGATCTTGTACTTTCACAAAGCCCTTAAAGCAGGTGAGCCAGCAATTCGCTGCA 540
Qy 537 AGTGCATCCACATGACAAAGCAAGAGTGGCAGAGCTTCAAGTTTGAAGAAACAG 596
    |||
Db 541 AGTGCATCCACATGACAAAGCAAGAGTGGCAGAGCTTCAAGTTTGAAGAAACAG 600
Qy 597 GAATCTCATATGATGATTTCTCCATGCTGACATACGGAAGAGTGTGTCCCTTG 656
    |||
Db 601 GAATCTCATATGATGATTTCTCCATGCTGACATACGGAAGAGTGTGTCCCTTG 660
Qy 657 GAGTCTCTGAAGTCTTCAACCTTGACCTTGGCTTCAAGCTGACGACGACATTC 716
    |||
Db 661 GAGTCTCTGAAGTCTTCAACCTTGACCTTGGCTTCAAGCTGACGACGACATTC 720
Qy 717 TAGATGAGCCCTTGGATAGAGGTGATGACATGCTTGTGCTGGCTCTTGGAGAAC 776
    |||
Db 721 TAGATGAGCCCTTGGATAGAGGTGATGACATGCTTGTGCTGGCTCTTGGAGAAC 780
Qy 777 TAAGAGATTTTACAGACGCTATATACAGAAAGTGTTCGAAATATCCAGAGACTATC 836
    |||
Db 781 TAAGAGATTTTACAGACGCTATATACAGAAAGTGTTCGAAATATCCAGAGACTATC 840
Qy 837 AACATGATCTTCCCAATCAATTTGGCTTCAAGAAATTTCCAGAGTACCTGATACAGAG 896
    |||
Db 841 AACATGATCTTCCCAATCAATTTGGCTTCAAGAAATTTCCAGAGTACCTGATACAGAG 900
Qy 897 GAAATGACACACTGAGAGCTAGTAAACAGGTTCTAAAGAAAG 939
    |||
Db 901 GAAATGACACACTGAGAGCTAGTAAACAGGTTCTAAAGAAAG 943

```

```

RESULT 7
US-10-097-340-150
; Sequence 150, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MOUNHAN
; APPLICANT: Manjula GANNANAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVARS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY

```

```

APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: US/10/097,340
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 150
LENGTH: 2129
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-150

Query Match      40.3%; Score 823; DB 39; Length 2129;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 823; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 AATCCAGCGTGGCGCTTGACGTAGGCGGCGCTGGCGGTGAGATGCTGACGCGCTGACT 176
DB 121 AATCCAGCGTGGCGCTTGACGTAGGCGGCGCTGGCGGTGAGATGCTGACGCGCTGACT 180
OY 177 CCATGCGAGGTATGAGAGGCTTACATCATCACCACAGGTTCTGCCAAGGCGA 236
DB 181 CCATGCGAGGTATGAGAGGCTTACATCATCACCACAGGTTCTGCCAAGGCGA 240
OY 237 GAATCTGCGGCGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 296
DB 241 GAATCTGCGGCGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
OY 297 TGAATCTGCGGCGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 356
DB 301 TGAATCTGCGGCGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
OY 357 CTATTTGTTGGGAGAACCAATTTATCATTTGATGATGATGATGATGATGATGATGATG 416
DB 361 CTATTTGTTGGGAGAACCAATTTATCATTTGATGATGATGATGATGATGATGATGATG 420
OY 417 ATACCAAGCCCCAGAGATGGGCGACTGGAAGATGATGACCAAGAAAGTGGAGCTTGA 476
DB 421 ATACCAAGCCCCAGAGATGGGCGACTGGAAGATGATGACCAAGAAAGTGGAGCTTGA 480
OY 477 AGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 536
DB 481 AGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
OY 537 AGCTGATTCACATGACAAAGCGCAAGTGGCGAGAGCTTCCAAATTTTGAAGAAAG 596
DB 541 AGCTGATTCACATGACAAAGCGCAAGTGGCGAGAGCTTCCAAATTTTGAAGAAAG 600
OY 597 GAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 656

```

```

DB 601 GAATCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
OY 657 GAGTCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 715
DB 661 GAGTCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
OY 717 TAGATGAGCGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 776
DB 721 TAGATGAGCGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780
OY 777 TAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 836
DB 781 TAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
OY 837 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 896
DB 841 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
OY 897 GAAATGACACCTGAGAGCTAGTAAACGAGCTTCTAAAGAAAG 939
DB 901 GAAATGACACCTGAGAGCTAGTAAACGAGCTTCTAAAGAAAG 943

RESULT 8
US-60-172-360-26034
Sequence 26034, Application US/60172360
GENERAL INFORMATION:
APPLICANT: Morris, Macdonald
APPLICANT: Lai, Precl
APPLICANT: Diep, Dinh
TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms
TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
FILE REFERENCE: GX-0007 P
CURRENT APPLICATION NUMBER: US/60/172,360
CURRENT FILING DATE: 1999-12-16
NUMBER OF SEQ ID NOS: 29838
SOFTWARE: PERL Program
SEQ ID NO 26034
LENGTH: 2101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 330917.13
NAME/KEY: unsure
LOCATION: 2098
OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-26034

Query Match      40.1%; Score 818; DB 61; Length 2101;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 939 GACCTGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 998
DB 1007 GACCTGTCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1056
OY 999 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1058
DB 1067 AAGAGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1126
OY 1059 CTACAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1118
DB 1127 CTACAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
OY 1119 TGTGTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1178
DB 1187 TGTGTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
OY 1179 CCAATGCCACTTGACCACTGAAGAAAGAAAGATTTGAGCTCAGATCTGTCAACA 1238

```

Db 1247 CCAATCCCACTTGAAACCAACGAGAGAGAAAGAGATTGGACTGATGCTGTCAACA 1306  
 Oy 1239 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGATCCC 1298  
 Db 1307 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGATCCC 1366  
 Oy 1299 CAGGCGAATATGATCAAGAGCTGAAATGCAAGCTTTAGAGACATGCTCCAGTGGCTTTG 1358  
 Db 1367 CAGGCGAATATGATCAAGAGCTGAAATGCAAGCTTTAGAGACATGCTCCAGTGGCTTTG 1426  
 Oy 1359 GAAAGTGGGGGAGATCCAGTTCCAGAGGAGGAGGTATGTTTCTCCAGTCTGGGCAA 1418  
 Db 1427 GAAAGTGGGGGAGATCCAGTTCCAGAGGAGGAGGTATGTTTCTCCAGTCTGGGCAA 1486  
 Oy 1419 AGAGAGTCTATGCGGAATTCCTGATACAGAGAAAGCTCCCACTTTTCTTTTATGAT 1478  
 Db 1487 AGAGAGTCTATGCGGAATTCCTGATACAGAGAAAGCTCCCACTTTTCTTTTATGAT 1546  
 Oy 1479 GGTTTTAAAGTCTCAGCTTCTCTATATATAGAAACAGCAGGTCTTGTACCTCCTTG 1538  
 Db 1547 GGTTTTAAAGTCTCAGCTTCTCTATATATAGAAACAGCAGGTCTTGTACCTCCTTG 1606  
 Oy 1539 GGTGATGTCTGGAATATGATGATGCTGAGAAAGCATTTTCTTTTGAACCTTAA 1598  
 Db 1607 GGTGATGTCTGGAATATGATGATGCTGAGAAAGCATTTTCTTTTGAACCTTAA 1666  
 Oy 1599 AGGTCTATTTTAAAGCAGACAGATCCACATTTTATACATGAGATCTTCTTTGT 1658  
 Db 1667 AGGTCTATTTTAAAGCAGACAGATCCACATTTTATACATGAGATCTTCTTTGT 1726  
 Oy 1659 GGTGAATACAGAGATTGATGCTGATCCCTTTAAAGAGTTTATGTCCCTGATCTGGCT 1718  
 Db 1727 GGTGAATACAGAGATTGATGCTGATCCCTTTAAAGAGTTTATGTCCCTGATCTGGCT 1785  
 Oy 1729 AAAATATCTAATTTCCAGATGCTTTTGTAGTGAAGTATTTGTGAGCCACATATT 1778  
 Db 1786 AAAATATCTAATTTCCAGATGCTTTTGTAGTGAAGTATTTGTGAGCCACATATT 1845  
 Oy 1779 GGGAGTCTAGATTGATGATGATGCGAGAAAGGCGCATTCCTGATGATCATTAAGTG 1838  
 Db 1846 GGGAGTCTAGATTGATGATGATGCGAGAAAGGCGCATTCCTGATGATCATTAAGTG 1905  
 Oy 1839 AACCAACTAGTCTCTCGAATTTCTACAGAGAGAGGGAATCAGACTAGAGAACTGTGA 1898  
 Db 1906 AACCAACTAGTCTCTCGAATTTCTACAGAGAGAGGGAATCAGACTAGAGAACTGTGA 1965  
 Oy 1899 CATGAGATCTGAAGACCAAAACCTTTGAAATTTGCGAGCTGCTCATGTGATTTATAT 1958  
 Db 1966 CATGAGATCTGAAGACCAAAACCTTTGAAATTTGCGAGCTGCTCATGTGATTTATAT 2025  
 Oy 1959 CACTGCTGCTTTTCTATGAGTTCAAAATCTATATTTTATTTGAAGTTAAATAAGAAA 2018  
 Db 2026 CACTGCTGCTTTTCTATGAGTTCAAAATCTATATTTTATTTGAAGTTAAATAAGAAA 2085  
 Oy 2019 AAATTTACAGA 2030  
 Db 2086 AAATTTACAGA 2097

RESULT 9  
 US-09-757-028-874  
 : Sequence 874, Application US/09757028  
 : GENERAL INFORMATION:  
 : APPLICANT: Rosen et al.  
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 : FILE REFERENCE: PM001  
 : CURRENT APPLICATION NUMBER: US/09/757,028  
 : CURRENT FILING DATE: 2001-01-09  
 : PRIOR APPLICATION NUMBER: 60/179,065  
 : PRIOR FILING DATE: 2000-01-31  
 : PRIOR APPLICATION NUMBER: 60/180,628  
 : PRIOR FILING DATE: 2000-02-04  
 : NUMBER OF SEQ ID NOS: 2660

: SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 874  
 : LENGTH: 1183  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: SITE  
 : LOCATION: (1172)  
 : OTHER INFORMATION: n equals a,t,g, or c  
 : NAME/KEY: SITE  
 : LOCATION: (1177)  
 : OTHER INFORMATION: n equals a,t,g, or c  
 : US-09-757-028-874

Query Match 38.1%; Score 778; DB 29; Length 1183;  
 Best Local Similarity 99.5%; Pred. No. 0;  
 Matches 1098; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Oy 939 GACCTGTGTCCTTGAACCTGCTTGAATGTCGAAAGTTTCATCCAGGCCCAACAG 1058  
 Db 35 GACCTGTGTCCTTGAACCTGCTTGAATGTCGAAAGTTTCATCCAGGCCCAACAG 154  
 Oy 999 AGAGTCTGTCTTGAACCTGCTTGAATGTCGAAAGTTTCATCCAGGCCCAACAG 1058  
 Db 95 AGAGTCTGTCTTGAACCTGCTTGAATGTCGAAAGTTTCATCCAGGCCCAACAG 154  
 Oy 1059 CTACAGCCACTCCCAATTAAGATGCCATACATGAAGCTGAGAACAGAAAGTTATAC 1118  
 Db 155 CTACAGCCACTCCCAATTAAGATGCCATACATGAAGCTGAGAACAGAAAGTTATAC 214  
 Oy 1119 TGTGTGACCTCTGTATGATGATCATCATTTGGGATGCGGAATGGGACGACATMAAT 1178  
 Db 215 TGTGTGACCTCTGTATGATGATCATCATTTGGGATGCGGAATGGGACGACATMAAT 274  
 Oy 1179 CCAATCCCACTGGAACCAAGCAAGAAAGAAAGAAAGTTGAGCTCAGATCTGTCAACA 1238  
 Db 275 CCAATCCCACTGGAACCAAGCAAGAAAGAAAGAAAGTTGAGCTCAGATCTGTCAACA 334  
 Oy 1239 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGATCCC 1298  
 Db 335 CCATAGAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGATCCC 394  
 Oy 1299 CAGGCGAATATGATCAAGAGCTGAAATGCAAGCTTTTAAAGACATGTCACATGGCTTTG 1358  
 Db 395 CAGGCGAATATGATCAAGAGCTGAAATGCAAGCTTTTAAAGACATGTCACATGGCTTTG 454  
 Oy 1359 GAAAGTGGTGGGATCCAGTTCAGAGAGGAGGGGTATGTTTGTCTCCAGTCTGGGCAA 1418  
 Db 455 GAAAGTGGTGGGATCCAGTTCAGAGAGGAGGGGTATGTTTGTCTCCAGTCTGGGCAA 514  
 Oy 1419 AGAGTCTATGCGGAATTCCTGATACAGAGAAAGCTCCCACTTTTCTTTTATGAT 1478  
 Db 515 AGAGTCTATGCGGAATTCCTGATACAGAGAAAGCTCCCACTTTTCTTTTATGAT 574  
 Oy 1479 GGTTTTAAAGTCTCAGCTTCTCTATATATAGAAACAGCAGGTCTTGTAGCTTGTG 1538  
 Db 575 GGTTTTAAAGTCTCAGCTTCTCTATATATAGAAACAGCAGGTCTTGTAGCTTGTG 634  
 Oy 1539 GGTGATGTCTGGAATATGATGATGCTGAGAAAGCATTTTCTTTTGAACCTTAA 1598  
 Db 635 GGTGATGTCTGGAATATGATGATGCTGAGAAAGCATTTTCTTTTGAACCTTAA 694  
 Oy 1599 AGGTCTATTTTAAAGCAGACAGATTCACATTTTATACATGAGATCTTCTTTGT 1658  
 Db 695 AGGTCTATTTTAAAGCAGACAGATTCACATTTTATACATGAGATCTTCTTTGT 754  
 Oy 1659 GGTGAATACAGAGATTGATGCTATCCCTTTAAAGAGTTTATGTCCTGACTGTGGCT 1718  
 Db 755 GGTGAATACAGAGATTGATGCTATCCCTTTAAAGAGTTTATGTCCTGACTGTGGCT 813  
 Oy 1719 AAATATATCAATTTCCAGATGCTTTTGTAGATGATGAGATTTTGTGAGCCACATATT 1778  
 Db 814 AAATATATCAATTTCCAGATGCTTTTGTAGATGATGAGATTTTGTGAGCCACATATT 873









QY 1601 GTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGG 1660  
|||||  
Db 441 GTCTATTATTAAGAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGG 382  
QY 1661 TGAATACAGAGATTCAGATCCCTTTAAAGAGTTTATGCTCCCTGACTGCTAA 1720  
|||||  
Db 381 TGAATACAGAGATTCAGATCCCTTTAAAGAGTTTATGCTCCCTGACTGCTAA 322  
QY 1721 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTACACCATATTGG 1780  
|||||  
Db 321 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATATTGG 262  
QY 1781 GAGTTCTAGATTTAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 1840  
|||||  
Db 261 GAGTTCTAGATTTAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 202  
QY 1841 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATGAGACTGAGGAAGCTGTGACA 1900  
|||||  
Db 201 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATGAGACTGAGGAAGCTGTGACA 142  
QY 1901 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGTGATTTATCA 1960  
|||||  
Db 141 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGTGATTTATCA 82  
QY 1961 CTGCTGCTTTCTATTGAGTACAAATCTATTTTATTGAGTTTAAATGAAGAAAA 2020  
|||||  
Db 81 CTGCTGCTTTCTATTGAGTACAAATCTATTTTATTGAGTTTAAATGAAGAAAA 22  
QY 2021 ATTTACAAGAAAAA 2041  
|||||  
Db 21 ATTTACAAGAAAAA 1  
RESULT 13  
US-09-888-8140/c  
; Sequence 8140, Application US/09898888  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
; FILE REFERENCE: 20411-748CON1  
; CURRENT APPLICATION NUMBER: US/09/898, 888  
; CURRENT FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 09/340, 623  
; PRIOR FILING DATE: 1999-06-28  
; NUMBER OF SEQ ID NOS: 45207  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8140  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-888-8140  
Query Match 37.1% Score 757; DB 33; Length 1041;  
Best Local Similarity 99.6% Pred. No. 0;  
Matches 1037; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1243 AGAAGTCAGAGTGTTCCTCCAGACTATACAAAGAACCTAAAGGAGGATCCCCAGG 1302  
|||||  
Db 801 AGAAGTCAGAGTGTTCCTCCAGACTATACAAAGAACCTAAAGGAGGATCCCCAGG 742  
QY 1303 GCAGAAATGATCAAGAGCTGAAATGACAGCTTTAAGACATGTCAGTGGCTTTGGAA 1362  
|||||  
Db 741 GCAGAAATGATCAAGAGCTGAAATGACAGCTTTAAGACATGTCAGTGGCTTTGGAA 682  
QY 1363 GGTGCTGGGATCCAGTTCAGAGAGGAGGGGATGTTGTTGCTCCAGTGGGCAAGAA 1422  
|||||  
Db 681 GGTGCTGGGATCCAGTTCAGAGAGGAGGGGATGTTGTTGCTCCAGTGGGCAAGAA 622  
QY 1423 GTGCTATGCGGAATTCCTCTCATAGCAGAAAGCTCCACCATTTCTTTTGAATGTTG 1482  
|||||  
Db 621 GTGCTATGCGGAATTCCTCTCATAGCAGAAAGCTCCACCATTTCTTTTGAATGTTG 562  
QY 1483 TTAAGTCTCAGCTTCTATTAATAGAACAGCAGGCTTGTGACGCTTGTGTGCTG 1542  
|||||  
Db 561 TTAAGTCTCAGCTTCTATTAATAGAACAGCAGGCTTGTGACGCTTGTGTGCTG 502  
QY 1543 ATGCTGCTGGAATGATGATGTTGAGAAAGCA--TTTTTTTTCTTTGACCTTTAAG 1600  
|||||  
Db 501 ATGCTGCTGGAATGATGATGTTGAGAAAGCAATTTTTTTTTCTTTGACCTTTAAG 442  
QY 1601 GTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGG 1660  
|||||  
Db 441 GTCTATTATTAAGCAGACAGATTCACATTTTATACATGAGGATCTCTTTGG 382  
QY 1661 TGAATACAGAGATTCAGATCCCTTTAAAGAGTTTATGCTCCCTGACTGCTAA 1720  
|||||  
Db 381 TGAATACAGAGATTCAGATCCCTTTAAAGAGTTTATGCTCCCTGACTGCTAA 322  
QY 1721 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATATTGG 1780  
|||||  
Db 321 AATTATCTAATTTCCAGATGCTTTTGTAGATGACTGAAGTATTTGTAGCCACATATTGG 262  
QY 1781 GAGTTCTAGATTTAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 1840  
|||||  
Db 261 GAGTTCTAGATTTAGTGAATGGCAGAAAGGCCATCTCCATGAGATGATTAAGTAA 202  
QY 1841 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATGAGACTGAGGAAGCTGTGACA 1900  
|||||  
Db 201 CCAAACTAGTTCTCGAATTTCTACAGAAAGGAGGAATGAGACTGAGGAAGCTGTGACA 142  
QY 1901 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGTGATTTATCA 1960  
|||||  
Db 141 TAGACTTTGAAGCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGTGATTTATCA 82  
QY 1961 CTGCTGCTTTCTATTGAGTACAAATCTATTTTATTGAGTTTAAATGAAGAAAA 2020  
|||||  
Db 81 CTGCTGCTTTCTATTGAGTACAAATCTATTTTATTGAGTTTAAATGAAGAAAA 22  
QY 2021 ATTTACAAGAAAAA 2041  
|||||  
Db 21 ATTTACAAGAAAAA 1  
RESULT 14  
US-09-888-8140/c  
; Sequence 8140, Application US/09898888  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
; FILE REFERENCE: 20411-748CON1  
; CURRENT APPLICATION NUMBER: US/09/898, 888  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/340, 623  
; PRIOR FILING DATE: 1999-06-28  
; PRIOR APPLICATION NUMBER: US 09/205, 070  
; PRIOR FILING DATE: 1998-12-03  
; NUMBER OF SEQ ID NOS: 45207

```

: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 8140
:
: LENGTH: 1041
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
US-09-898-888A-8140
.
```

Query Match	37.18;	Score 757;	DB 33;	length 1041;
Best Local Similarity	99.68;	Pred. NO. 0;		
Matches 1037; Conservative	0;	Mismatches 2;	Indels 2;	Gaps 1;

Oy 1003 GTCTGTTCTTACACCTGCTCTTGAATGCTGCAAGTTTCAATCCAGGGCCACAAGCTTC 1062  
 Db 1041 GTCTGTTCTTACACCTGCTCTTGAATGCTGCAAGTTTCAATCCAGGGCCACAAGCTTC 982  
 Oy 1063 AGCCACTTCGAATTAAGATGCCCTTACAATGAAGCTGAGAACAAGAAAGTTATACCTGTG 1122  
 Db 981 AGCCACTTCGAATTAAGATGCCCTTACAATGAAGCTGAGAACAAGAAAGTTATACCTGTG 922  
 Oy 1123 TGACCTCTGTGATGGAATCATCATTTGGGGATCCGCAATGGGACGGCACAATTAATCCAA 1182  
 Db 921 TGACCTCTGTGATGGAATCATCATTTGGGGATCCGCAATGGGACGGCACAATTAATCCAA 862  
 Oy 1183 ATCCCACTTGAACCAACTGGAAGAAAGAAAGATTTGAGCTCAGATGCTGTCAACCAT 1242  
 Db 861 ATCCCACTTGAACCAACTGGAAGAAAGAAAGATTTGAGCTCAGATGCTGTCAACCAT 802  
 Oy 1243 AGAAAGTCAGAGTGTTCCTCCAGACTATACAAAGACCTTAAAGGAAAGGATCCCAAG 1302  
 Db 801 AGAAAGTCAGAGTGTTCCTCCAGACTATACAAAGACCTTAAAGGAAAGGATCCCAAG 742  
 Oy 1303 GCACAATGATCAAGAGCGGAAGTCAAGGCTTAAAGACATGTCCAGAGGCTTTGGAAA 1362  
 Db 741 GCACAATGATCAAGAGCGGAAGTCAAGGCTTAAAGACATGTCCAGAGGCTTTGGAAA 682  
 Oy 1363 GGTGTGGGGATCCAGTTCAAGAGGAGGAGGTATGTTGTCTCCAGTCTGGCAAGAA 1422  
 Db 681 GGTGTGGGGATCCAGTTCAAGAGGAGGAGGTATGTTGTCTCCAGTCTGGCAAGAA 622  
 Oy 1423 GTGCTATGCGGAATCTCTGATAGCAGAAAAGCTCCACCATTTCTTTTGATGTGTT 1482  
 Db 621 GTGCTATGCGGAATCTCTGATAGCAGAAAAGCTCCACCATTTCTTTTGATGTGTT 562  
 Oy 1483 TTAAAGTCTCAGTCTCTATATATAGAAAAGCAGAGTGTGTCAAGCTCTGTGGGTG 1542  
 Db 561 TTAAAGTCTCAGTCTCTATATATAGAAAAGCAGAGTGTGTCAAGCTCTGTGGGTG 502  
 Oy 1543 ATGTGCTCGAATGATGTAGTTCAGAGAAACA - - TTTTCTTTTCTTTTGAACCTTAAAG 1600  
 Db 501 ATGTGCTCGAATGATGTAGTTCAGAGAAACATTTTCTTTTCTTTGAACCTTAAAG 442  
 Oy 1601 GTTCTATTTTAAAGAGCAGACAGATTCACATTTTATACATAGAGATCTCTTTGGG 1660  
 Db 441 GTTCTATTTTAAAGAGCAGACAGATTCACATTTTATACATAGAGATCTCTTTGGG 382  
 Oy 1661 TGAATACAGAGATTTGACTGCATCCCTTAAAGAAAGTTTATGTGCCGACTGTGGCTAA 1720  
 Db 381 TGAATACAGAGATTTGACTGCATCCCTTAAAGAAAGTTTATGTGCCGACTGTGGCTAA 322  
 Oy 1721 AATTATCTAATTTTCCAGATGCTTTTGTAGATGACTGGAAGTATTTGTGAGCCACATATTGG 1780  
 Db 321 AATTATCTAATTTTCCAGATGCTTTTGTAGATGACTGGAAGTATTTGTGAGCCACATATTGG 262  
 Oy 1781 GAGTTCAGATTTGAGAGTATGGCAGAGAAAGGCGCATCTCCATTTGAGATGATTTAAAGTAA 1840  
 Db 261 GAGTTCAGATTTGAGAGTATGGCAGAGAAAGGCGCATCTCCATTTGAGATGATTTAAAGTAA 202  
 Oy 1841 CCAAACTAGTTCGCAATTTCTACAGAGAAAGAGGAAATCAGACTGAGAGAGCTGTGACA 1900  
 Db 201 CCAAACTAGTTCGCAATTTCTACAGAGAAAGAGGAAATCAGACTGAGAGAGCTGTGACA 142  
 Oy 1901 TAGGACTTGAGAGCAAAAGCTTTGAAATTTGAGAGTGTCTCATGTGTGAGTTATTATCA 1960  
 Db 142 TAGGACTTGAGAGCAAAAGCTTTGAAATTTGAGAGTGTCTCATGTGTGAGTTATTATCA 142

[illegible]

RESULT 15  
PCT-US02-18947-1692

```
Sequence 1692, Application PC/TUS0218947
```

FILE REFERENCE: 9301-175-228

```

: CURRENT APPLICATION NUMBER: PCT/US02/18947
: CURRENT FILING DATE: 2002-06-14

```

PRIOR FILING DATE: 2002-05-14  
NUMBER OF SEQ TO NOS: 3699  
PRIOR APPLICATION NUMBER: 60/380,110

```

      SEQ ID NO 1692
      LENGTH: 1749

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
;

```

```

      PUBLICATION INFORMATION:
      DATABASE ACCESSION NUMBER: NM_017646
      DATABASE ENTRY DATE: 2001-05-18

```

PCT-US02-18947-1692

Query Match	36.88;	Score 751;	DB 1;	Length 1749;
Best Local Similarity	99.68;	Pred. No. 0;		
Number of Sequences	1001			

	conservative	/	medicine	/	index	1,	cups	1
by	1001	GAGTCTGTTCTTGACCGTCGTGAATCGTCGCAAAAGTTTCATCCAGGGCCACAAGCCT	1060					

Db 706 GACTCTGTCCTTGAACCTGCTCTTGAATCTGCAAGTTTCATCCAGGGCCACAGCCT 765

QY 1061 ACAGCCACTCCATTAAGATGCCATACCAATGGAAGCTGAGAACAGAGAGATTTCACCTG 1120

D<sub>b</sub> 766 ACAGCCACTCCATAAAGATGCCATACAAATGAAGCTGAGACCAAGAGAGTTATCACCCTG 825

Qy 1121 TGTGACCTCTGTATCGATCATCTATTGGGATCCGGAATGGGACGGCACCATAAAATTC 1180

[illegible]

886 AATCCCACTTGAAACCACTGAAGAAAAGAGAAGAGATTGGACTCAGATGCTGTCAACACC 945

QY 1241 ATAGAACTCAGAGTGTTCCTCCAGACTATATACAAAGACCTAAAGGAGGATCCCA 1300

Db 946 ATAGAAGTCAAGTGTTCCTCCCGACCAATAACAAGAAGCTAAAGAGAAGGATCCCCA 1005

QY 1301 GGGCAGATGATCAAGAGCGTGAATGCGCGCTTTAAGAGACATGTCAGTGGCCTTTGGA 1360  
|||||

[illegible]

1066 AAGCTGCGGATTCACGAGGAGGGGTATGTTGTCCTCCAGTCGSGGCAAG 1125

Oy 1421 GAGTGCATATCGGAATTTCTGCATAGCAGAAAAGCTCCACACATTTTCTTTTGATGTGG 1480

Db 1126 GAGTGCATGCCGATTTCTTGCATAGCAGAAAAAGCTCCCAACCATTTTCTTTTGATGTGG 1185

QY 1481 TTTTAAAGTCTCAGCTTCTCTATATATAGAAACACAGGCTTTGTCA6CTCCmTGTGTGGC 1540

DB 1186 TTTTAAAGTCACCGTCTCTATAATAGAAACACAGGCTTTGTCAAGCTCCCTTGTGTGGC 1245

```
Db 1246 TGATGTGCTCGAATGATGATTCAGAGAACATTTTCTTTCTTGAACCTTAAG 1305
QY 1601 GTTCTATTATTAAGACGACAGATTCACATTTTATACATGAGATCTTCTTGTG 1660
Db 1306 GTTCTATTATTAAGACGACAGATTCACATTTTATACATGAGATCTTCTTGTG 1365
QY 1661 TGAATACGAGATGACATGCCCTTAAAGAGTTTATGTCCCTGACTGCTGCTAA 1720
Db 1366 TGAATACGAGATGACATGCCCTT -AAAGAGTTTATGTCCCTGACTGCTGCTAA 1424
QY 1721 AATTATCTAATTCACAGATGCTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTG 1780
Db 1425 AATTATCTAATTCACAGATGCTTTGTAGATGACTGAAGTATTTGTGAGCCACATATTG 1484
QY 1781 GAGTTCTAGATTGAGTGAATGCGAGAAAGGCCATCTCATGAGATGATTAAGTGA 1840
Db 1485 GAGTTCTAGATTGAGTGAATGCGAGAAAGGCCATCTCATGAGATGATTAAGTGA 1544
QY 1841 CCAACTAGATTCTCGAATTTCTACAGAGAGAGGAAATCAGACTGAGAGAGCTGTGACA 1900
Db 1545 CCAACTAGATTCTCGAATTTCTACAGAGAGAGGAAATCAGACTGAGAGAGCTGTGACA 1604
QY 1901 TAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGTCATGTGAGTTATTATCA 1960
Db 1605 TAGGACTTGAAGACCAAGACTTTGAAATTTGCGAGCTGTCATGTGAGTTATTATCA 1664
QY 1961 CTGCTGCTTCTATGAGTACAAATCTATATTTTATGAGTTTAAATAAGAAAAA 2020
Db 1665 CTGCTGCTTCTATGAGTACAAATCTATATTTTATGAGTTTAAATAAGAAAAA 1724
QY 2021 ATTTA 2025
Db 1725 ATTTA 1729
```

Search completed: April 21, 2003, 23:43:40  
Job time : 5197 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:34:21 ; Search time 85 Seconds  
(Without alignments)  
7363.850 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 2041

Sequence: 1 CTCGCAATAGATGCGTCG.....TTTACAAAGAAAAA 2041

Scoring table: OLIGO\_NUC  
Gapop 60.0, Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/5A.COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/5B.COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	23	1.1 87350	3 US-08-781-891-79	Sequence 79, Appl
2	23	1.1 87543	4 US-09-791-211-3	Sequence 3, Appl
3	21	1.0 98844	4 US-09-791-211-10	Sequence 10, Appl
4	19	0.9 1314	4 US-09-134-001C-2505	Sequence 2505, Ap
5	19	0.9 1659	4 US-09-152-060-24	Sequence 24, Appl
6	19	0.9 49136	4 US-09-422-869-1	Sequence 1, Appl
7	18	0.9 293	1 US-08-222-177A-49	Sequence 49, Appl
8	18	0.9 618	4 US-09-289-349-2	Sequence 2, Appl
9	18	0.9 657	4 US-09-134-001C-2195	Sequence 2195, Ap
10	18	0.9 1319	2 US-08-484-938B-17	Sequence 17, Appl
11	18	0.9 1319	2 US-08-484-158B-17	Sequence 17, Appl
12	18	0.9 1319	2 US-08-484-596A-17	Sequence 17, Appl
13	18	0.9 1319	2 US-08-480-150A-17	Sequence 17, Appl
14	18	0.9 1319	3 US-08-458-731-17	Sequence 17, Appl
15	18	0.9 1319	3 US-08-149-223A-17	Sequence 17, Appl
16	18	0.9 1761	3 US-09-043-830-2	Sequence 2, Appl
17	18	0.9 2100	1 US-08-485-718-10	Sequence 10, Appl
18	18	0.9 2100	2 US-08-484-530-56	Sequence 56, Appl
19	18	0.9 2100	2 US-08-827-618A-56	Sequence 56, Appl
20	18	0.9 2100	3 US-08-483-952A-56	Sequence 56, Appl
21	18	0.9 2100	4 US-08-476-501-56	Sequence 56, Appl
22	18	0.9 2100	6 5475086-5	Patent No. 5475086
23	18	0.9 2249	2 US-08-494-624-1	Sequence 1, Appl
24	18	0.9 2370	1 US-08-117-907-1	Sequence 1, Appl
25	18	0.9 2370	3 US-08-453-040-1	Sequence 1, Appl
26	18	0.9 2756	1 US-08-187-793-1	Sequence 1, Appl
27	18	0.9 3318	1 US-08-187-793-3	Sequence 3, Appl

C 28	18	0.9 5973	4 US-09-245-041-4	Sequence 4, Appl
C 29	18	0.9 6623	2 US-08-687-080-68	Sequence 68, Appl
C 30	18	0.9 21234	4 US-09-810-671-3	Sequence 3, Appl
C 31	18	0.9 32042	4 US-09-245-281-44	Sequence 44, Appl
C 32	18	0.9 168575	4 US-09-426-290-1	Sequence 1, Appl
C 33	18	0.9 176373	3 US-09-128-155-17	Sequence 17, Appl
C 34	17	0.8 40	4 US-09-306-290-31	Sequence 31, Appl
C 35	17	0.8 260	1 US-08-594-031-45	Sequence 45, Appl
C 36	17	0.8 260	1 US-08-594-031-137	Sequence 137, Appl
C 37	17	0.8 260	1 US-08-594-031-144	Sequence 144, App
C 38	17	0.8 272	2 US-08-454-557C-100	Sequence 100, App
C 39	17	0.8 272	2 US-08-340-426D-100	Sequence 100, App
C 40	17	0.8 272	2 US-08-450-673C-100	Sequence 100, App
C 41	17	0.8 272	5 PCT-US95-17111A-100	Sequence 100, App
C 42	17	0.8 324	4 US-08-651-155B-22	Sequence 22, Appl
C 43	17	0.8 452	3 US-09-033-055A-2	Sequence 2, Appl
C 44	17	0.8 456	4 US-08-978-289-3	Sequence 3, Appl
C 45	17	0.8 604	4 US-09-370-838-175	Sequence 175, App

## ALIGNMENTS

## RESULT 1

US-08-781-891-79  
Sequence 79, Application US/08781891

Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,891

FILING DATE: 27-DEC-1996

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: No. 6090620tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 87350 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-781-891-79

Query Match: 1.1%; Score 23; DB 3; Length 87350;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2019 AATTTACAGAAAAA 2041  
|||||

Db 51760 AAATTACAGAAAAA 51782

## RESULT 2

US-09-791-211-3

Sequence 3, Application US/09791211

Patent No. 6448080

GENERAL INFORMATION:

APPLICANT: Donna T. Ward

APPLICANT: Andrew T. Walt

TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION

FILE REFERENCE: RTS-0205

CURRENT APPLICATION NUMBER: US/09/791,211

CURRENT FILING DATE: 2001-02-23

NUMBER OF SEQ ID NOS: 90

SEQ ID NO 3

LENGTH: 87543

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: 7421

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 7427

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 11609

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 12605

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 12742

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29370

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29422

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29979

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29980

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 29981

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 30136

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 30140

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 31205

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 31206

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 31592

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 33095

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 33160

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 34072

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 36816

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 39020

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 42164

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 42459

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 46808

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 46823

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 46826

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 47291

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 52786

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 52787

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 53384

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 54684

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 59215

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 59235

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 59242

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 63290

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 66614

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68660

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68697

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68718

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68733

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 68739

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 69785

OTHER INFORMATION: unknown

NAME/KEY: unsure  
LOCATION: 79134  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79198  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 86336  
OTHER INFORMATION: unknown  
OTHER INFORMATION: unknown  
US-09-791-211-3

Query Match 1.18; Score 23; DB 4; Length 87543;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTACAGAAAAA 2041  
|||||  
DB 51953 AAATTACAGAAAAA 51975

RESULT 3  
US-09-791-211-10/c  
Sequence 10, Application US/09791211  
Patent No. 6448080  
GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
FILE REFERENCE: RTS-0205  
CURRENT APPLICATION NUMBER: US/09/791.211  
CURRENT FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 90  
SEQ ID NO 10  
LENGTH: 98844  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 24962  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 64383  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65468  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65469  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65470  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 65471  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 87130  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 89049  
OTHER INFORMATION: unknown  
OTHER INFORMATION: unknown  
US-09-791-211-10

Query Match 1.0%; Score 21; DB 4; Length 98844;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2019 AAATTACAGAAAAA 2039  
|||||  
DB 36142 AAATTACAGAAAAA 36122

RESULT 4  
US-09-134-001C-2505  
Sequence 2505, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134.001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/035,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2505  
LENGTH: 1314  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2505

Query Match 0.9%; Score 19; DB 4; Length 1314;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1595 TTAAGGTTCTATTATTA 1613  
|||||  
DB 1274 TTAAGGTTCTATTATTA 1292

RESULT 5  
US-09-152-060-24/c  
Sequence 24, Application US/09152060  
Patent No. 6448230  
GENERAL INFORMATION:  
APPLICANT: Rosen et al  
TITLE OF INVENTION: 28 Human Secreted Proteins  
FILE REFERENCE: P2003P1.US  
CURRENT APPLICATION NUMBER: US/09/152.060  
CURRENT FILING DATE: 1998-09-11  
EARLIER APPLICATION NUMBER: PCT/US98/04858  
EARLIER FILING DATE: 1998-03-12  
EARLIER APPLICATION NUMBER: 60/040,762  
EARLIER FILING DATE: 1997-03-14  
EARLIER APPLICATION NUMBER: 60/040,710  
EARLIER FILING DATE: 1997-03-14  
EARLIER APPLICATION NUMBER: 60/050,934  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/048,100  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/048,357  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/048,189  
EARLIER FILING DATE: 1997-05-30  
EARLIER APPLICATION NUMBER: 60/057,765  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/068,368  
EARLIER FILING DATE: 1997-12-19  
NUMBER OF SEQ ID NOS: 118  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 24  
LENGTH: 1659  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-152-060-24

Query Match 0.9%; Score 19; DB 4; Length 1659;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 388 TGAATCTGCTGGAA 406  
|||||  
Db 311 TGAATCTGCTGGAA 293

## RESULT 6

US-09-422-869-1/C  
Sequence 1, Application US/09422869  
Patent No. 6235481  
GENERAL INFORMATION:  
APPLICANT: POLONSKY, KENNETH S.  
APPLICANT: HORIKAWA, YUKIO  
APPLICANT: ODA, NAOHISA  
APPLICANT: COX, NANCY J.  
APPLICANT: SREENAN, SEAMUS  
APPLICANT: ZHOU, YUN-PING  
APPLICANT: OTANI, KENICHI  
APPLICANT: HANIS, CRAIG L.  
APPLICANT: BELL, GRAEME I.  
TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES  
FILE REFERENCE: ARCD:307  
CURRENT APPLICATION NUMBER: US/09/422,869  
CURRENT FILING DATE: 1999-10-21  
EARLIER APPLICATION NUMBER: 60/134,175  
EARLIER FILING DATE: 1999-05-13  
NUMBER OF SEQ. ID NOS: 30  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 49136  
TYPE: DNA  
ORGANISM: Human  
US-09-422-869-1

Query Match 0.9%; Score 19; DB 4; Length 49136;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1091 GAAGCTGAGAACAGAGAA 1109  
|||||  
Db 35847 GAAGCTGAGAACAGAGAA 35829

## RESULT 7

US-08-222-177A-49/C  
Sequence 49, Application US/08222177A  
Patent No. 5382979  
GENERAL INFORMATION:  
APPLICANT: Weber, James L.  
TITLE OF INVENTION: LENGTH POLYMORPHISMS IN  
TITLE OF INVENTION: (dc-da)n.(dc-dt)n SEQUENCES AND METHODS OF USING SAME  
NUMBER OF SEQUENCES: 460  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dewitt Ross & Stevens, S.C.  
STREET: 8000 Excelstor Drive, Suite 401  
CITY: Madison  
STATE: Wisconsin  
COUNTRY: USA  
ZIP: 53717-1914  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,177A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/341,562  
FILING DATE: 21-APR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Sara, Charles S.

REGISTRATION NUMBER: 30,492  
REFERENCE/DOCKET NUMBER: 09865.601  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 831-2100  
TELEFAX: (608) 831-2106  
TELEX:

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 293 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

INDIVIDUAL ISOLATE: Caucasian

TISSUE TYPE: Blood

IMMEDIATE SOURCE:

CLONE: Mcd122

POSITION IN GENOME:

CHROMOSOME/SEGMENT: 5q

FEATURE:

NAME/KEY: repeat\_region

LOCATION: 41..74

OTHER INFORMATION: /rpt\_type="tandem"

OTHER INFORMATION: /rpt\_family="(dc-da)n.(dc-dt)n"

OTHER INFORMATION: /citation="(12)"

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 4..23

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence=EXPERIMENTAL

OTHER INFORMATION: /standard\_name="PCR primer"

OTHER INFORMATION: /citation="(11)"

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1..293

IDENTIFICATION METHOD: experimental

OTHER INFORMATION: /evidence=EXPERIMENTAL

OTHER INFORMATION: /standard\_name="Only one strand sequenced"

PUBLICATION INFORMATION:

AUTHORS: Weber, J. L.

AUTHORS: Polymeropoulos, M. H.

AUTHORS: May, P. E.

AUTHORS: Kwitek, A. E.

AUTHORS: Xiao, H.

AUTHORS: McPherson, J. D.

AUTHORS: Masmuth, J. J.

TITLE: Mapping of human chromosome 5 microsatellite

TITLE: polymorphisms

JOURNAL: Genomics

DATE: 1991

PUBLICATION INFORMATION:

AUTHORS: Weber, James L.

AUTHORS: May, Paula E.

TITLE: Abundant Class of Human DNA Polymorphisms

TITLE: Which Can Be Typed Using the Polymerase Chain

TITLE: Reaction

JOURNAL: Am. J. Hum. Genet.

VOLUME: 44

PAGES: 388-396

US-08-222-177A-49  
Query Match 0.9%; Score 18; DB 1; Length 293;  
Best Local Similarity 100.0%; Pred. No. 46;

Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2019 AATTACAGAAAAA 2036  
|||||  
DB 197 AATTACAGAAAAA 180

## RESULT 8

US-09-289-349-2/C  
Sequence 2, Application US/09289349  
Patent No. 627/574  
GENERAL INFORMATION:  
APPLICANT: Walker, Michael, G.  
APPLICANT: Volkmut, Wayne  
APPLICANT: Klinger, Tod, M.  
APPLICANT: Azimzal, Yalda  
APPLICANT: Yue, Henry  
TITLE OF INVENTION: GENES ASSOCIATED WITH DISEASES OF THE KIDNEY  
FILE REFERENCE: PB-0010 US  
CURRENT APPLICATION NUMBER: US/09/289,349  
CURRENT FILING DATE: 1999-04-09  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PERL Program  
SEQ ID NO 2  
LENGTH: 618  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: 3534377C71  
US-09-289-349-2

Query Match 0.9%; Score 18; DB 4; Length 618;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2024 TACAGAAAAA 2041  
|||||  
DB 578 TACAGAAAAA 561

## RESULT 9

US-09-134-001C-2195  
Sequence 2195, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2195  
LENGTH: 657  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2195

Query Match 0.9%; Score 18; DB 4; Length 657;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 AATAAGAAAAATTTA 2025  
|||||  
DB 210 AATAAGAAAAATTTA 227

RESULT 10  
US-08-484-993B-17

Sequence 17, Application US/08484993B  
Patent No. 5837497

GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immuncontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,993B  
FILING DATE: 09-NOV-1993

CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.

REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448

TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 1319 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Felis domesticus

DEVELOPMENTAL STAGE: Juvenile

HAPLOTYPE: Diploidy

TISSUE TYPE: Ovary

CELL TYPE: oocyte

FEATURE:  
NAME/KEY: CDS  
LOCATION: 26..1297

US-08-484-993B-17

Query Match 0.9%; Score 18; DB 2; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1702 TGTCCTGACTGGCTA 1719  
|||||  
DB 1206 TGTCCTGACTGGCTA 1223

## RESULT 11

US-08-484-158B-17  
Sequence 17, Application US/08484158B  
Patent No. 5976545  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.



APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Pharmaceutical Compositions for  
TITLE OF INVENTION: Immunoccontraception  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
ADDRESSEE: Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,158B  
FILING DATE: 07-JUNE-95  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,223  
FILING DATE: 09-NOV-93  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-93  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-92  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 32794  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ. ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1319 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Felis domesticus  
DEVELOPMENTAL STAGE: Juvenile  
HAPLOTYPE: Diploidy  
TISSUE TYPE: Ovary  
CELL TYPE: Oocyte  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 26..1297  
US-08-484-158B-17

Query Match 0.9%; Score 18; DB 2; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1702 TGTCCCTGACTCTGGCTA 1719  
|||||  
Db 1206 TGTCCCTGACTCTGGCTA 1223

RESULT 12  
US-08-484-596A-17  
Sequence 17, Application US/08484596A  
Patent No. 5981228  
GENERAL INFORMATION:

APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunoccontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,596A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/149,223  
FILING DATE: 11-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ. ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1319 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Felis domesticus  
DEVELOPMENTAL STAGE: Juvenile  
HAPLOTYPE: Diploidy  
TISSUE TYPE: Ovary  
CELL TYPE: Oocyte  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 26..1297  
US-08-484-596A-17

Query Match 0.9%; Score 18; DB 2; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1702 TGTCCCTGACTCTGGCTA 1719  
|||||  
Db 1206 TGTCCCTGACTCTGGCTA 1223

RESULT 13  
US-08-480-150A-17  
Sequence 17, Application US/08480150A  
Patent No. 5989550  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunoccontraception

NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,150A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/149,223  
FILING DATE: 09-NOV-1993  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1319 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Felis domesticus  
DEVELOPMENTAL STAGE: Juvenile  
HAPLOTYPE: Diploidy  
TISSUE TYPE: Ovary  
CELL TYPE: Oocyte  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 26..1297  
US-08-480-150A-17  
Query Match 0.9%; Score 18; DB 2; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1702 TGTCCTGACTGTGCTA 1719  
|||||  
Db 1206 TGTCCTGACTGTGCTA 1223  
RESULT 14  
US-08-458-731-17  
Sequence 17, Application US/08458731  
Patent No. 6001599  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,731  
FILING DATE: 09-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1319 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Felis domesticus  
DEVELOPMENTAL STAGE: Juvenile  
HAPLOTYPE: Diploidy  
TISSUE TYPE: Ovary  
CELL TYPE: Oocyte  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 26..1297  
US-08-458-731-17  
Query Match 0.9%; Score 18; DB 3; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1702 TGTCCTGACTGTGCTA 1719  
|||||  
Db 1206 TGTCCTGACTGTGCTA 1223  
RESULT 15  
US-08-149-223A-17  
Sequence 17, Application US/08149223A  
Patent No. 6027727  
GENERAL INFORMATION:  
APPLICANT: Harris Ph.D., Jeffrey D.  
APPLICANT: Hsu, Kuang T.  
APPLICANT: Podolski, Joseph S.  
TITLE OF INVENTION: Materials and Methods for Immunocontraception  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois

COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,223A  
FILING DATE: 09-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/012,990  
FILING DATE: 29-JAN-1993  
PRIOR APPLICATION DATA: 07/973,341  
FILING DATE: 09-NOV-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Clough, David W.  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 31745  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6653  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1319 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Felis domesticus  
DEVELOPMENTAL STAGE: Juvenile  
HAPLOTYPE: Diploidy  
TISSUE TYPE: Ovary  
CELL TYPE: Oocyte  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 26..1297  
US-08-149-223A-17

Query Match 0.98; Score 18; DB 3; Length 1319;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Caps 0;  
QY 1702 TGTCCCTGACTCTGGCTA 1719  
Db 1206 TGTCCCTGACTCTGGCTA 1223

Search completed: April 21, 2003, 23:50:13  
Job time : 393 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 19:21:06 ; Search time 480 Seconds

(Without alignments)  
9575.685 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 2041

Sequence: 1 CTGCCATAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: OLIGO-NUC  
Gap 60.0, Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

```
N.Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2041	100.0	2041	20	AA336073
2	896	43.9	2130	24	ABN59708
3	473	23.2	717	22	ABA06500
4	473	23.2	717	22	ABA06598
5	473	23.2	717	22	AA541268
6	241	11.8	355	24	ABL68710
7	175	8.6	300	21	AAA00911
8	145	7.1	457	22	ABA42839
9	145	7.1	457	22	ABA53262

C 10	145	7.1	457	22	AAK01519
C 11	145	7.1	457	22	AAK26966
C 12	145	7.1	457	22	AAI11574
C 13	145	7.1	457	22	AAI32862
C 14	138	6.8	138	22	ABA47966
C 15	138	6.8	138	22	ABA55854
C 16	138	6.8	138	22	AAK14265
C 17	138	6.8	138	22	AAK39988
C 18	138	6.8	138	22	AAI20787
C 19	138	6.8	138	22	AAI46014
C 20	60	2.9	60	24	ABN34729
C 21	59	2.9	582	23	ABV49413
C 22	40	2.0	545	23	ABV19646
C 23	23	1.1	2001	23	AA588985
C 24	23	1.1	4557	23	AA579149
C 25	23	1.1	6881	22	AAI92597
C 26	23	1.1	7313	23	AA581397
C 27	23	1.1	7488	22	AAK87243
C 28	23	1.1	7488	23	AAK87437
C 29	23	1.1	11087	23	AA574637
C 30	23	1.1	11378	23	AA587336
C 31	23	1.1	13234	23	AA582685
C 32	23	1.1	15879	23	AA584109
C 33	23	1.1	33780	22	AAH24652
C 34	22	1.1	87350	18	AAK83003
C 35	22	1.1	885	22	AAK78227
C 36	22	1.1	1118	22	AAI36412
C 37	22	1.1	1398	22	AAI36415
C 38	22	1.1	1398	22	AAK66112
C 39	22	1.1	6225	23	ABL10424
C 40	22	1.1	6225	23	ABL06640
C 41	22	1.1	6399	23	ABL17248
C 42	22	1.1	19802	23	ABL12950
C 43	22	1.1	33376	22	AAK66113
C 44	22	1.1	33376	24	ABN85746
C 45	22	1.1	33357	24	ABN85746

#### ALIGNMENTS

RESULT 1  
ID AAX36073  
AAX36073 standard: DNA; 2041 BP.

AC AAX36073:  
15-JUL-1999 (first entry)

DE Human homologue of gro-1, referred to as hgro-1.

KW gro-1 operon: gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;  
hap-1 gene; cancer; aging; longevity; tumour formation;  
physiological clock; ss.

KW Homo sapiens.  
WO9910482-A1.

PN 04-MAR-1999.  
20-AUG-1998; 98MO-CA00803.  
25-AUG-1997; 97CA-2210251.

PR (UYMC-) UNIV MCGILL.  
Barnes T, Hekimi S, Lakowski B, Lemieux J;  
WPI: 1999-190615/16.

PT Molecular identity of the gro-1 gene - useful for cancer diagnosis  
and/or prognosis, and where compounds affecting encoded proteins are

Human brain expres  
Human bone marrow  
Probe #1507 for ge  
Probe #1548 used t  
Human breast cell  
Human foetal liver  
Human brain expres  
Human bone marrow  
Probe #10720 for g  
Probe #14700 used  
Human spliced tran  
Human prostate exp  
Human prostate exp  
DNA encoding novel  
DNA encoding novel  
Human polynucleoti  
DNA encoding novel  
Human immune/haema  
Human immune/haema  
DNA encoding novel  
DNA encoding novel  
DNA encoding novel  
DNA encoding novel  
Nucleotide sequenc  
Human WPN genomic  
Human secreted pro  
Human immune/haema  
Human musculoskele  
Human immune/haema  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Human immune/haema  
Human transporter

PT useful for enhancing longevity of a host and inhibiting tumour  
PT formation  
XX  
PS Claim 8; Fig 8; 93pp; English.

XX The present sequence represents the human homologue of gro-1, and  
CC is referred to as hgro-1. The specification describes the five genes  
CC of the Caenorhabditis elegans gro-1 operon (AXX36071). The operon  
CC contains the gro-1 gene (AXX36072), the gop-1 gene (AXX36074), the gop-2  
CC gene (AXX36075), the gop-2 gene (AXX36075), and the hap-1 gene  
CC (AXX36077). The gro-1 gene can be used in a method for the diagnosis  
CC and/or prognosis of cancer in a patient. Transgenic mice containing a  
CC gene knock-out of a murine gene homologue of the gro-1 gene are useful  
CC as models of aging and cancer. The proteins encoded by the genes are  
CC useful for identifying compounds that affect the enzymatic activity  
CC of these proteins. In order to enhance longevity of a host and inhibit  
CC tumour formation, the gro-1 gene, together with the gop-1, gop-2,  
CC gop-3 and hap-1 genes enables study of a physiological clock.

XX Sequence 2041 BP: 589 A; 421 C; 502 G; 529 T; 0 other;

Query Match 100.0%; Score 2041; DB 20; Length 2041;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 1 CTGCCAATAAGATGGCGCGCGGCGCTGACAGCAGCTTCCTGTGGGAGTGGGCTCAG 60
DB 1 CTGCCAATAAGATGGCGCGCGGCGCTGACAGCAGCTTCCTGTGGGAGTGGGCTCAG 60
OY 61 GGGCTGCAAGGAGACCTTACTCTTGTAGTATTCGGGGCCACGGGCAACGCAATC 120
DB 61 GGGCTGCAAGGAGACCTTACTCTTGTAGTATTCGGGGCCACGGGCAACGCAATC 120
OY 121 CAGGCTGGGCTTGCAGTACGCGGCGCTGGGCGGTGAGATGTCACGCTGACTCCAT 180
DB 121 CAGGCTGGGCTTGCAGTACGCGGCGCTGGGCGGTGAGATGTCACGCTGACTCCAT 180
OY 181 GCAGTCTATGAGGCTTAGACATCATCACCACAAGGCTTTCGCCAAGAGCAGAGAT 240
DB 181 GCAGTCTATGAGGCTTAGACATCATCACCACAAGGCTTTCGCCAAGAGCAGAGAT 240
OY 241 CTGCGGCGACACATGATGAGCTTGTGATCTCTTGTGACCAATTCACAGTGTGTA 300
DB 241 CTGCGGCGACACATGATGAGCTTGTGATCTCTTGTGACCAATTCACAGTGTGTA 300
OY 301 CTTCAGAAATAGAGCAACTGCTCTGATGTAAGATATATTTGCCCGAGCAAAATTCCTAT 360
DB 301 CTTCAGAAATAGAGCAACTGCTCTGATGTAAGATATATTTGCCCGAGCAAAATTCCTAT 360
OY 361 TGTGTGGGAGGAAACCAATTAATGATGATCTGCTGTGAAAGTTCTGTCAATAC 420
DB 361 TGTGTGGGAGGAAACCAATTAATGATGATCTGCTGTGAAAGTTCTGTCAATAC 420
OY 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTGAATGACCCGAAAGTGGAGCTTGAAGAAG 480
DB 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTGAATGACCCGAAAGTGGAGCTTGAAGAAG 480
OY 481 GATGCTCTTGTACTTCAAAAGCGCTAAGCCAGGTGAGACCCAGAAATGGCTCCAGCT 540
DB 481 GATGCTCTTGTACTTCAAAAGCGCTAAGCCAGGTGAGACCCAGAAATGGCTCCAGCT 540
OY 541 GCATCCACATGACAAACGCAAAAGTGCGAGAGCTTGCAAGTTTGAAGAAACAGAGAT 600
DB 541 GCATCCACATGACAAACGCAAAAGTGCGAGAGCTTGCAAGTTTGAAGAAACAGAGAT 600
OY 601 CTCTCATAGTGAATTTCTCATGTCACATACGAGAGAGGTGGTCTCCCTTGGAG 660
DB 601 CTCTCATAGTGAATTTCTCATGTCACATACGAGAGAGGTGGTCTCCCTTGGAG 660
OY 661 TCTCTGAAGTTCTCTAACCTTGGATCTTGGCTTCAATGCTGACAGCAGGTTCTAGA 720
DB 661 TCTCTGAAGTTCTCTAACCTTGGATCTTGGCTTCAATGCTGACAGCAGGTTCTAGA 720
```

```
OY 721 TGAGCGCTTGATTAAGGCGTGGATGACATGCTTGCCTGCGGCTCTTGGAGCACTAAC 780
DB 721 TGAGCGCTTGATTAAGGCGTGGATGACATGCTTGCCTGCGGCTCTTGGAGCACTAAC 780
OY 781 AGATTTTACAGAGCGCTATATATCAGAGAAATGTTTGGAAAAATAGCCAGGACTATACA 840
DB 781 AGATTTTACAGAGCGCTATATATCAGAGAAATGTTTGGAAAAATAGCCAGGACTATACA 840
OY 841 TGGTATCTTCAATCAATTTGGCTTCAAGAAATTTACAGATACCTGATCACTGAGGAAA 900
DB 841 TGGTATCTTCAATCAATTTGGCTTCAAGAAATTTACAGATACCTGATCACTGAGGAAA 900
OY 901 ATGCACCTGAGACACTGTATACCAAGCTTCTTAAGAAAGACCTGCTGCTTCCCTCC 960
DB 901 ATGCACCTGAGACACTGTATACCAAGCTTCTTAAGAAAGACCTGCTGCTTCCCTCC 960
OY 961 TGTCTATGCTTGAAGTATCTGATGCTGCAAGTGGAGAGAGTCTTCTTGAACCTGC 1020
DB 961 TGTCTATGCTTGAAGTATCTGATGCTGCAAGTGGAGAGAGTCTTCTTGAACCTGC 1020
OY 1021 TCTTGAATGCTGCAAGTCTTCAATCCAGGGCCACAAAGCTTACAGCCACTCAATAAGAT 1080
DB 1021 TCTTGAATGCTGCAAGTCTTCAATCCAGGGCCACAAAGCTTACAGCCACTCAATAAGAT 1080
OY 1081 GCCATACATGAACTGAGAACAGAAATATATCACTGTGACCTCTGTATGAAAT 1140
DB 1081 GCCATACATGAACTGAGAACAGAAATATATCACTGTGACCTCTGTATGAAAT 1140
OY 1141 CATCATTTGGGAGTGGCAATGAGGCGAGGCGCATATAAATCCAAATCCACTTGAACCACT 1200
DB 1141 CATCATTTGGGAGTGGCAATGAGGCGAGGCGCATATAAATCCAAATCCACTTGAACCACT 1200
OY 1201 GAAGAAAGAAAGAAATTTGAGCTCAGCTCAGCAACCATAGAAAGTACAGAGTGTTC 1260
DB 1201 GAAGAAAGAAAGAAATTTGAGCTCAGCTCAGCAACCATAGAAAGTACAGAGTGTTC 1260
OY 1261 CCCAGACTATACAAAGAAACCTTAAGGAGGATCCCGGCGAGCAATGATATAAGAGCT 1320
DB 1261 CCCAGACTATACAAAGAAACCTTAAGGAGGATCCCGGCGAGCAATGATATAAGAGCT 1320
OY 1321 GAAATGCAGCGTTTAAAGACATGTCAGTGGCTTTGGAAGAGTGGTGGGATCCAGTT 1380
DB 1321 GAAATGCAGCGTTTAAAGACATGTCAGTGGCTTTGGAAGAGTGGTGGGATCCAGTT 1380
OY 1381 CAGAGAGGAGGATGATGTTGCTCCCACTCTGGGCAAGAGAGTCTATGCGAATTCCTC 1440
DB 1381 CAGAGAGGAGGATGATGTTGCTCCCACTCTGGGCAAGAGAGTCTATGCGAATTCCTC 1440
OY 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGAGTGTGTAAGTCTCACGTTCTC 1500
DB 1441 TGCATAGCAGAAAGCTCCACCATTTCTTTGATGAGTGTGTAAGTCTCACGTTCTC 1500
OY 1501 TATATATAGAAACAGCAGGCTTGTCACTCTTGTGCTGATGCTGTGAAATGATG 1560
DB 1501 TATATATAGAAACAGCAGGCTTGTCACTCTTGTGCTGATGCTGTGAAATGATG 1560
OY 1561 TAGTTAGAGAAACATTTTTTTTCTTTGAACCTTAAAGTTCTATTAATAAGCAGC 1620
DB 1561 TAGTTAGAGAAACATTTTTTTTCTTTGAACCTTAAAGTTCTATTAATAAGCAGC 1620
OY 1621 ACAGATTCACATTTTATACATGAGAGATCTTCTTGTGTAATACAGAGATGACCTGC 1680
DB 1621 ACAGATTCACATTTTATACATGAGAGATCTTCTTGTGTAATACAGAGATGACCTGC 1680
OY 1681 ATCCCTTTAAAGAAAGTTTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
DB 1681 ATCCCTTTAAAGAAAGTTTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
OY 1741 CTTTGTAGATGATGAGATTTTGTGAGCCACATATTTGGAGTCTTGAATTTGAGTGA 1800
DB 1741 CTTTGTAGATGATGAGATTTTGTGAGCCACATATTTGGAGTCTTGAATTTGAGTGA 1800
OY 1801 TGGCAGGAAAGGCGCATCTCCATTTGATGATTAAGTGAACCAACTAGTCTCGGAATT 1860
```

|||||  
Db 1801 TCGAGGAGAAAGGCCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTCTCGAATT 1860  
Qy 1861 CTACAGAGAGAGAGGGAATGAGTGAAGAGCTGTGACATAGAGACTTTGAAGACCAAGA 1920  
Db 1861 CTACAGAGAGAGAGGGAATGAGTGAAGAGCTGTGACATAGAGACTTTGAAGACCAAGA 1920  
Qy 1921 CTTTGAATTTGCGAGCTGCTCATGTGTGAGTATTTATCATGTGCTTCTTTTGTAGT 1980  
Db 1921 CTTTGAATTTGCGAGCTGCTCATGTGTGAGTATTTATCATGTGCTGCTTCTTTTGTAGT 1980  
Qy 1981 TACAAATCTATATTTTATTTGAGTAAATTAAGAAAAAATTTACAGAAAAA 2040  
Db 1981 TACAAATCTATATTTTATTTGAGTAAATTAAGAAAAAATTTACAGAAAAA 2040  
Qy 2041 A 2041  
Db 2041 A 2041  
RESULT 2  
ABN59708  
ID ABN59708 standard; cDNA: 2130 BP.  
XX  
AC ABN59708:  
XX  
DT 28-JUN-2002 (first entry)  
XX  
DE Novel human coding sequence SEQ ID NO: 119.  
XX  
KW Homo sapiens.  
OS  
PN MO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX  
PF 10-SEP-2001; 2001WO-US26015.  
XX  
PR 11-SEP-2000; 2000US-0659671.  
PA (HYSE-) HYSEQ INC.  
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F,  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
DR WPI: 2002-292408/33.  
DR P-PSDB: ABB97295.  
PT An isolated polynucleotide for treating diseases associated with its  
XX encoded polypeptide such as cancer and multiple sclerosis -  
XX  
PS Claim 1; SEQ ID NO 119; 509pp: English.  
XX  
CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a coding sequence of the  
XX invention.  
XX  
SQ Sequence 2130 BP; 620 A; 439 C; 522 G; 549 T; 0 other;

Query Match

43.9%; Score 896; DB 24; Length 2130;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 896; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 44 GTGGGCACTGGGCTCAGGGGGCTTCAACGGACCTTACTCTTGTAGTATCTCTGGGGCC 103  
Db 56 GTGGGCACTGGGCTCAGGGGGCTTCAACGGACCTTACTCTTGTAGTATCTCTGGGGCC 115  
Qy 104 ACGGAGACCGGCAATCCACGCTGCGTGTGAGTAGAGCCAGCGGCTCGGGGTGAGATC 163  
Db 116 ACGGAGACCGGCAATCCACGCTGCGTGTGAGTAGAGCCAGCGGCTCGGGGTGAGATC 175  
Qy 164 GTCAAGCTGACTCCATGACAGTCTATGAAAGCCCTAGACATCATCACCAAGGTTCT 223  
Db 176 GTCAAGCTGACTCCATGACAGTCTATGAAAGCCCTAGACATCATCACCAAGGTTCT 235  
Qy 224 GCCCAAGAGCAGAGAAATCTGCGGACACATGATCAGCTTTGTGATCTCTTGTGACC 283  
Db 236 GCCCAAGAGCAGAGAAATCTGCGGACACATGATCAGCTTTGTGATCTCTTGTGACC 295  
Qy 284 AATTACACAGTGTGGAGTCTGAGAAATAGACAACTGCTGATTTGAAGATATATTTGCC 343  
Db 296 AATTACACAGTGTGGAGTCTGAGAAATAGACAACTGCTGATTTGAAGATATATTTGCC 355  
Qy 344 CGAGACAAATCTCTATTTGTGTGGAGGAGAACCAATTATTCATTGATGATCTCTGCTGG 403  
Db 356 CGAGACAAATCTCTATTTGTGTGGAGGAGAACCAATTATTCATTGATGATCTCTGCTGG 415  
Qy 404 AAAGTTCTTGTCAATACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCGAAAA 463  
Db 416 AAAGTTCTTGTCAATACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCGAAAA 475  
Qy 464 GTGAGCTTGAAGAGAGATGCTGTCTACTTACAAACCCCTTAAGCAGCTGGAGCCA 523  
Db 476 GTGAGCTTGAAGAGAGATGCTGTCTACTTACAAACCCCTTAAGCAGCTGGAGCCA 535  
Qy 524 GAAATGCGTCCCAAGCTCATCCACATGACAAAGCGAGGCGGAGAGCTTCCAAAGTT 583  
Db 536 GAAATGCGTCCCAAGCTCATCCACATGACAAAGCGAGGCGGAGAGCTTCCAAAGTT 595  
Qy 584 TTGAGAGAAACAGAGAAATCTCTCATAGTGAATTTCTCCATGCTCAACATACGGAGAAAGT 643  
Db 596 TTGAGAGAAACAGAGAAATCTCTCATAGTGAATTTCTCCATGCTCAACATACGGAGAAAGT 655  
Qy 644 GGTGTCCTCCCTGGAGGCTCTGAGTCTGAGTCTCAACCTTCAACCTTGGCTTCATAGCT 703  
Db 656 GGTGTCCTCCCTGGAGGCTCTGAGTCTGAGTCTCAACCTTCAACCTTGGCTTCATAGCT 715  
Qy 704 GACCAGCAGTTCTAGATGAGCGCTTGGATGAAGAGGATGATGACATGCTGCTGGG 763  
Db 716 GACCAGCAGTTCTAGATGAGCGCTTGGATGAAGAGGATGATGACATGCTGCTGGG 775  
Qy 764 CTTTGGAGAGACTAGAGATTTTCACAGACGCTATTAATCAAGAAATGTTCCGAAAT 823  
Db 776 CTTTGGAGAGACTAGAGATTTTCACAGACGCTATTAATCAAGAAATGTTCCGAAAT 835  
Qy 824 AGCCAGACTATCAACATGATGATCTCCCAATCAATTTGGCTCAAGAAATTCACAGATAC 883  
Db 836 AGCCAGACTATCAACATGATGATCTCCCAATCAATTTGGCTCAAGAAATTCACAGATAC 895  
Qy 884 CTGATCACTAGAGGAAATGCACTGAGACTAGTAACAGCTTCTTAAGAAAGG 939  
Db 896 CTGATCACTAGAGGAAATGCACTGAGACTAGTAACAGCTTCTTAAGAAAGG 951  
RESULT 3  
ABA06500  
ID ABA06500 standard; cDNA: 717 BP.  
XX  
AC ABA06500:  
XX  
DT 10-JAN-2002 (first entry)  
XX  
DE Human cDNA SEQ ID NO: 166.

XX	Human: gene therapy; neural disorder; immune system disorder;	PR	14-SEP-2000;	2000US-233065P.
KM	muscular disorder; reproductive disorder; gastrointestinal disorder;	PR	21-SEP-2000;	2000US-234223P.
KM	pulmonary disorder; cardiovascular disorder; renal disorder;	PR	21-SEP-2000;	2000US-234274P.
KM	proliferative disorder; inflammation; ss.	PR	25-SEP-2000;	2000US-234997P.
XX		PR	25-SEP-2000;	2000US-234998P.
OS	Homo sapiens.	PR	26-SEP-2000;	2000US-235484P.
XX		PR	27-SEP-2000;	2000US-235834P.
PN	WO200154474-A2.	PR	27-SEP-2000;	2000US-235836P.
XX		PR	29-SEP-2000;	2000US-236327P.
XX		PR	29-SEP-2000;	2000US-236367P.
PD	02-AUG-2001.	PR	29-SEP-2000;	2000US-236368P.
XX		PR	29-SEP-2000;	2000US-236369P.
PF	17-JAN-2001; 2001WO-US01349.	PR	29-SEP-2000;	2000US-236370P.
XX		PR	02-OCT-2000;	2000US-236802P.
PR	31-JAN-2000; 2000US-179065P.	PR	02-OCT-2000;	2000US-237037P.
PR	04-FEB-2000; 2000US-160628P.	PR	02-OCT-2000;	2000US-237038P.
PR	24-FEB-2000; 2000US-184664P.	PR	02-OCT-2000;	2000US-237039P.
PR	02-MAR-2000; 2000US-186350P.	PR	02-OCT-2000;	2000US-237040P.
PR	16-MAR-2000; 2000US-189874P.	PR	13-OCT-2000;	2000US-239935P.
PR	17-MAR-2000; 2000US-190076P.	PR	13-OCT-2000;	2000US-239937P.
PR	18-APR-2000; 2000US-198123P.	PR	20-OCT-2000;	2000US-240960P.
PR	19-MAY-2000; 2000US-205515P.	PR	20-OCT-2000;	2000US-241221P.
PR	07-JUN-2000; 2000US-209467P.	PR	20-OCT-2000;	2000US-241222P.
PR	28-JUN-2000; 2000US-214886P.	PR	20-OCT-2000;	2000US-241785P.
PR	30-JUN-2000; 2000US-215135P.	PR	20-OCT-2000;	2000US-241786P.
PR	07-JUL-2000; 2000US-216647P.	PR	20-OCT-2000;	2000US-241787P.
PR	11-JUL-2000; 2000US-216880P.	PR	20-OCT-2000;	2000US-241808P.
PR	11-JUL-2000; 2000US-217487P.	PR	20-OCT-2000;	2000US-241809P.
PR	11-JUL-2000; 2000US-217496P.	PR	20-OCT-2000;	2000US-241826P.
PR	14-JUL-2000; 2000US-218290P.	PR	01-NOV-2000;	2000US-244617P.
PR	26-JUL-2000; 2000US-220963P.	PR	08-NOV-2000;	2000US-246474P.
PR	26-JUL-2000; 2000US-220964P.	PR	08-NOV-2000;	2000US-246475P.
PR	14-AUG-2000; 2000US-224518P.	PR	08-NOV-2000;	2000US-246476P.
PR	14-AUG-2000; 2000US-224519P.	PR	08-NOV-2000;	2000US-246477P.
PR	14-AUG-2000; 2000US-225213P.	PR	08-NOV-2000;	2000US-246478P.
PR	14-AUG-2000; 2000US-225214P.	PR	08-NOV-2000;	2000US-246523P.
PR	14-AUG-2000; 2000US-225266P.	PR	08-NOV-2000;	2000US-246525P.
PR	14-AUG-2000; 2000US-225267P.	PR	08-NOV-2000;	2000US-246526P.
PR	14-AUG-2000; 2000US-225268P.	PR	08-NOV-2000;	2000US-246527P.
PR	14-AUG-2000; 2000US-225270P.	PR	08-NOV-2000;	2000US-246528P.
PR	14-AUG-2000; 2000US-225447P.	PR	08-NOV-2000;	2000US-246532P.
PR	14-AUG-2000; 2000US-225757P.	PR	08-NOV-2000;	2000US-246609P.
PR	14-AUG-2000; 2000US-225758P.	PR	08-NOV-2000;	2000US-246610P.
PR	14-AUG-2000; 2000US-225759P.	PR	08-NOV-2000;	2000US-246611P.
PR	18-AUG-2000; 2000US-226279P.	PR	08-NOV-2000;	2000US-246613P.
PR	22-AUG-2000; 2000US-226681P.	PR	17-NOV-2000;	2000US-249207P.
PR	22-AUG-2000; 2000US-226861P.	PR	17-NOV-2000;	2000US-249208P.
PR	22-AUG-2000; 2000US-227182P.	PR	17-NOV-2000;	2000US-249209P.
PR	23-AUG-2000; 2000US-227009P.	PR	17-NOV-2000;	2000US-249210P.
PR	30-AUG-2000; 2000US-228924P.	PR	17-NOV-2000;	2000US-249211P.
PR	01-SEP-2000; 2000US-229287P.	PR	17-NOV-2000;	2000US-249212P.
PR	01-SEP-2000; 2000US-229343P.	PR	17-NOV-2000;	2000US-249213P.
PR	01-SEP-2000; 2000US-229344P.	PR	17-NOV-2000;	2000US-249214P.
PR	01-SEP-2000; 2000US-229345P.	PR	17-NOV-2000;	2000US-249215P.
PR	05-SEP-2000; 2000US-229509P.	PR	17-NOV-2000;	2000US-249216P.
PR	05-SEP-2000; 2000US-229513P.	PR	17-NOV-2000;	2000US-249217P.
PR	06-SEP-2000; 2000US-230437P.	PR	17-NOV-2000;	2000US-249218P.
PR	06-SEP-2000; 2000US-230438P.	PR	17-NOV-2000;	2000US-249244P.
PR	08-SEP-2000; 2000US-231242P.	PR	17-NOV-2000;	2000US-249245P.
PR	08-SEP-2000; 2000US-231243P.	PR	17-NOV-2000;	2000US-249246P.
PR	08-SEP-2000; 2000US-231244P.	PR	17-NOV-2000;	2000US-249255P.
PR	08-SEP-2000; 2000US-231413P.	PR	17-NOV-2000;	2000US-249297P.
PR	08-SEP-2000; 2000US-231414P.	PR	17-NOV-2000;	2000US-249299P.
PR	08-SEP-2000; 2000US-232080P.	PR	17-NOV-2000;	2000US-249300P.
PR	08-SEP-2000; 2000US-232081P.	PR	01-DEC-2000;	2000US-250160P.
PR	12-SEP-2000; 2000US-231968P.	PR	01-DEC-2000;	2000US-250391P.
PR	14-SEP-2000; 2000US-232397P.	PR	05-DEC-2000;	2000US-251030P.
PR	14-SEP-2000; 2000US-232398P.	PR	05-DEC-2000;	2000US-251988P.
PR	14-SEP-2000; 2000US-232399P.	PR	05-DEC-2000;	2000US-256719P.
PR	14-SEP-2000; 2000US-232400P.	PR	06-DEC-2000;	2000US-251479P.
PR	14-SEP-2000; 2000US-232401P.	PR	08-DEC-2000;	2000US-251856P.
PR	14-SEP-2000; 2000US-233063P.	PR	08-DEC-2000;	2000US-251868P.
PR	14-SEP-2000; 2000US-233064P.	PR	08-DEC-2000;	2000US-251869P.

PR 08-DEC-2000; 2000US-251989P.  
PR 08-DEC-2000; 2000US-251990P.  
PR 11-DEC-2000; 2000US-254097P.  
PR 05-JAN-2001; 2001US-259678P.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI: 2001-476161/51.  
DR P-PSDB; ABB10278.  
XX  
PT Isolated nucleic acid molecule encoding an inflammation-associated  
PT polypeptide is used in preventing, treating or ameliorating a medical  
PT condition  
XX  
XX Claim 1; SEQ ID NO: 166; 859pp + Sequence Listing; English.  
XX  
XX The present invention provides human CDNAs, proteins and related genomic  
CC DNAs. These can be used in the treatment of neural, immune system,  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
CC renal and proliferative disorders and inflammation. The present sequence  
CC is a cDNA of the invention.  
XX  
XX

Sequence 717 BP; 184 A; 176 C; 187 G; 162 T; 8 other;

Query Match 23.2%; Score 473; DB 22; Length 717;

Best Local Similarity 99.7%; Pred. No. 1.3e-213; Mismatches 2; Indels 0; Gaps 0;

Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 44 GTGGGAGTGGGCTGACGGGCTGCAAGGACCTGCTGTAGTGAATTCGCGGGCC 103  
DB 44 GTGGGAGTGGGCTGACGGGCTGCAAGGACCTGCTGTAGTGAATTCGCGGGCC 103  
QY 104 ACGGGACCGGCAATCCACGCTGGCTTGCAGCTAGGCGGCTGCGGCTGAGATC 163  
DB 104 ACGGGACCGGCAATCCACGCTGGCTTGCAGCTAGGCGGCTGCGGCTGAGATC 163  
QY 164 GTGAGCGCTGACTCCATGAGGTCTATGAAGGCTAGACATCATCAACAAGTTTCT 223  
DB 164 GTGAGCGCTGACTCCATGAGGTCTATGAAGGCTAGACATCATCAACAAGTTTCT 223  
QY 224 GCCCAAGACAGAGATTCGCCGCGCACATGATCAGCTTGTGATCTCTTGTGACC 283  
DB 224 GCCCAAGACAGAGATTCGCCGCGCACATGATCAGCTTGTGATCTCTTGTGACC 283  
QY 284 AATTACACAGTGTGACTTCAAGAAATAGAGCAACTGCTGTGATGAATATATTGGC 343  
DB 284 AATTACACAGTGTGACTTCAAGAAATAGAGCAACTGCTGTGATGAATATATTGGC 343  
QY 344 CGAGACAAATTCCTATTGTTGTGGAGAGAACCAATTATTACTTAATCTCGCTCTGG 403  
DB 344 CGAGACAAATTCCTATTGTTGTGGAGAGAACCAATTATTACTTAATCTCGCTCTGG 403  
QY 404 AAAGTCTTGTCAATACAGAGCCCGAGAGATGGGACCTAGAGAAATGATGACCAAAA 463  
DB 404 AAAGTCTTGTCAATACAGAGCCCGAGAGATGGGACCTAGAGAAATGATGACCAAAA 463  
QY 464 GTGAGACTTGAAAGAGAGATGCTTGTACTTACAAAAGCCTTAAGCAGGTGACCCA 523  
DB 464 GTGAGACTTGAAAGAGAGATGCTTGTACTTACAAAAGCCTTAAGCAGGTGACCCA 523  
QY 524 GAAATGGCTGCCAAGTGCATGCATGACAAAAGCAAGTGGCCAGGAGCTTGCAAGTT 563  
DB 524 GAAATGGCTGCCAAGTGCATGCATGACAAAAGCAAGTGGCCAGGAGCTTGCAAGTT 563  
QY 584 TTGGAAGAACAAGATCTCTATAGTAATTTCT 618  
DB 584 TTGGAAGAACAAGATCTCTATAGTAATTTCT 618

RESULT 4  
ABA06698

ID ABA06698 standard; CDNA; 717 BP.  
XX  
AC ABA06698;  
XX  
DT 10-JAN-2002 (first entry)  
XX  
DE Human CDNA SEQ ID NO: 364.  
XX  
XX Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200154474-A2.  
XX  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01349.  
XX  
XX 31-JAN-2000; 2000US-179065P.  
XX  
XX 04-FEB-2000; 2000US-180628P.  
XX  
XX 24-FEB-2000; 2000US-184664P.  
XX  
XX 02-MAR-2000; 2000US-186350P.  
XX  
XX 16-MAR-2000; 2000US-189874P.  
XX  
XX 17-MAR-2000; 2000US-190076P.  
XX  
XX 18-APR-2000; 2000US-198123P.  
XX  
XX 19-MAY-2000; 2000US-205515P.  
XX  
XX 07-JUN-2000; 2000US-209467P.  
XX  
XX 28-JUN-2000; 2000US-214866P.  
XX  
XX 30-JUN-2000; 2000US-215135P.  
XX  
XX 07-JUL-2000; 2000US-216647P.  
XX  
XX 07-JUL-2000; 2000US-216880P.  
XX  
XX 11-JUL-2000; 2000US-217487P.  
XX  
XX 11-JUL-2000; 2000US-217496P.  
XX  
XX 14-JUL-2000; 2000US-218290P.  
XX  
XX 26-JUL-2000; 2000US-220963P.  
XX  
XX 26-JUL-2000; 2000US-220964P.  
XX  
XX 14-AUG-2000; 2000US-224518P.  
XX  
XX 14-AUG-2000; 2000US-224519P.  
XX  
XX 14-AUG-2000; 2000US-225213P.  
XX  
XX 14-AUG-2000; 2000US-225214P.  
XX  
XX 14-AUG-2000; 2000US-225266P.  
XX  
XX 14-AUG-2000; 2000US-225267P.  
XX  
XX 14-AUG-2000; 2000US-225268P.  
XX  
XX 14-AUG-2000; 2000US-225270P.  
XX  
XX 14-AUG-2000; 2000US-225447P.  
XX  
XX 14-AUG-2000; 2000US-225757P.  
XX  
XX 14-AUG-2000; 2000US-225758P.  
XX  
XX 14-AUG-2000; 2000US-225759P.  
XX  
XX 18-AUG-2000; 2000US-226279P.  
XX  
XX 22-AUG-2000; 2000US-226681P.  
XX  
XX 22-AUG-2000; 2000US-226686P.  
XX  
XX 22-AUG-2000; 2000US-227182P.  
XX  
XX 23-AUG-2000; 2000US-227009P.  
XX  
XX 30-AUG-2000; 2000US-228924P.  
XX  
XX 01-SEP-2000; 2000US-229287P.  
XX  
XX 01-SEP-2000; 2000US-229343P.  
XX  
XX 01-SEP-2000; 2000US-229344P.  
XX  
XX 01-SEP-2000; 2000US-229345P.  
XX  
XX 05-SEP-2000; 2000US-229509P.  
XX  
XX 05-SEP-2000; 2000US-229513P.  
XX  
XX 06-SEP-2000; 2000US-230437P.  
XX  
XX 06-SEP-2000; 2000US-230438P.  
XX  
XX 08-SEP-2000; 2000US-231242P.  
XX  
XX 08-SEP-2000; 2000US-231243P.  
XX  
XX 08-SEP-2000; 2000US-231244P.  
XX  
XX 08-SEP-2000; 2000US-231413P.  
XX  
XX 08-SEP-2000; 2000US-231414P.  
XX  
XX 08-SEP-2000; 2000US-232080P.  
XX  
XX 08-SEP-2000; 2000US-232081P.  
XX  
XX 12-SEP-2000; 2000US-231968P.



Query Match	Best Local Similarity	Score	DB	Length
Matches 573; Conservative	0; Mismatches 2; Indels	0; Gaps	0;	0;
Sequence 717 BP; 184 A; 177 C; 187 G; 162 T; 7 other:				
0Y 44 GTGGGACAGGGGCTCAGGGGCTCGAACGACGACCTACCTCTGTAGTATTCGCGGGCC	103			
DB 44 GTGGGACAGGGGCTCAGGGGCTCGAACGACGACCTACCTCTGTAGTATTCGCGGGCC	103			
0Y 104 ACGGGACACGGGAAATCCACGCGCTGGCTGACAGTACGACGCGCTCGCGGTGAGATC	163			
DB 104 ACGGGACACGGGAAATCCACGCGCTGGCTGACAGTACGACGCGCTCGCGGTGAGATC	163			
0Y 164 GTCACGCTGACTCATGACAGTCTATGTAGAGGCTTAGACATATCACCAAGCTTCT	223			
DB 164 GTCACGCTGACTCATGACAGTCTATGTAGAGGCTTAGACATATCACCAAGCTTCT	223			
0Y 224 GCCCAAGACAGAGATCTGCCGGACCCAGATCAGCTTGTGGAATCTCTTGAGCC	283			
DB 224 GCCCAAGACAGAGATCTGCCGGACCCAGATCAGCTTGTGGAATCTCTTGAGCC	283			
0Y 284 AATTACACAGTGGTGGACTTCAGAATATGAGACACTGCTCTGATTGAAGATATTTGCC	343			
DB 284 AATTACACAGTGGTGGACTTCAGAATATGAGACACTGCTCTGATTGAAGATATTTGCC	343			
0Y 344 CGAGCAAAATTCCTATTTGTTGGGAGAGACCAATTATTACATTGAATCTCTGCTGG	403			
DB 344 CGAGCAAAATTCCTATTTGTTGGGAGAGACCAATTATTACATTGAATCTCTGCTGG	403			
0Y 404 AAAGTCTTGTCAATATACCAAGCCCGGAGAGATGGGACACTGAGAAAGTGATACCGGAAA	463			
DB 404 AAAGTCTTGTCAATATACCAAGCCCGGAGAGATGGGACACTGAGAAAGTGATACCGGAAA	463			
0Y 464 GTGAGCTTGAAGAGAGATGCTTGTACTTCAAAAGCGCTAAGCAGGTGACCA	523			
DB 464 GTGAGCTTGAAGAGAGATGCTTGTACTTCAAAAGCGCTAAGCAGGTGACCA	523			
0Y 524 GAAATGGCTGCCAAGCTGCATCAGATGACAAAGCGAAAGTGCGCAGAGCTTGAAGTT	583			
DB 524 GAAATGGCTGCCAAGCTGCATCAGATGACAAAGCGAAAGTGCGCAGAGCTTGAAGTT	583			

OY 584 TTTGAGAAACGAGATCTCTCATAGTGATTTCT 618  
|||||  
DB 584 TTTGAGAAACGAGATCTCTCATAGTGATTTCT 618  
RESULT 5  
AAS41268  
ID AAS41268 standard; cDNA; 717 BP.  
AC AAS41268;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE CDNA encoding novel human enzyme polypeptide #484.  
XX  
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
KW anti arthritic; nephrotropic; anticoagulant; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO20015301-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01239.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227709.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI: 2001-465566/50.

DR P-PSDB; AAU23398.

PT Novel polypeptides and polynucleotides useful for diagnosing,  
 PT preventing, treating neural, immune system, muscular, reproductive,  
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
 PT diseases

PS Claim 4; SEQ ID NO 494; 1180bp; English.

XX The present invention relates to the isolation of novel human enzyme  
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
 CC encoding them. The enzyme polypeptides of the invention may comprise the  
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
 CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
 CC blood-related disorders (e.g. hemophilia), reproductive disorders  
 CC (e.g. infertility) and infectious disorders (e.g. influenza). The  
 CC polynucleotides of the invention can also be used in gene therapy.  
 CC AAS40785-AAS41684 represent cDNA sequences encoding for the novel human  
 CC enzyme polypeptides of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 717 BP; 184 A; 177 C; 187 G; 162 T; 7 other:

Query Match 23 2%; Score 473; DB 22; Length 717;

Best Local Similarity 99.7%; Pred. No. 1.3e-213;

Matches 573; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 44 GTGGGAGTGGGCTCAGGGGCTGCAAGGACCTACCTCTTGTAGTATTCGGGGC 103  
 DB 44 GTGGGAGTGGGCTCAGGGGCTGCAAGGACCTACCTCTTGTAGTATTCGGGGC 103  
 QY 104 ACGGGACCGGCAAAATTCACGCTGGCGTTGACCTAGGCGCGCTGGCGTGACATC 163  
 DB 104 ACGGGACCGGCAAAATTCACGCTGGCGTTGACCTAGGCGCGCTGGCGTGACATC 163  
 QY 164 GTGAGGCTGACTCCATGCAAGGTCTATGAAGGCTAGACTATCATCAACAGAGTTTCT 223  
 DB 164 GTGAGGCTGACTCCATGCAAGGTCTATGAAGGCTAGACTATCATCAACAGAGTTTCT 223

QY 224 GCCAAGACGAGAGATCTGCCGGCACCACATGATACGCTTGTGATCCTCTTGACC 283  
 DB 224 GCCAAGACGAGAGATCTGCCGGCACCACATGATACGCTTGTGATCCTCTTGACC 283  
 QY 284 AATTACACAGTGTGACCTTGAGAAATAGACCACTGCTGTGATGAAGATATTTGGC 343  
 DB 284 AATTACACAGTGTGACCTTGAGAAATAGACCACTGCTGTGATGAAGATATTTGGC 343  
 QY 344 CGAGACAAATTCCTATTGTTGTGGGAGAACCAATTATTCATTGATATCTGCTGCG 403  
 DB 344 CGAGACAAATTCCTATTGTTGTGGGAGAACCAATTATTCATTGATATCTGCTGCG 403  
 QY 404 AAATTTCTTGTCATTACCAAGCCCGAGAGATGGGCACTGGAAGATGATTGACCGAAA 463  
 DB 404 AAATTTCTTGTCATTACCAAGCCCGAGAGATGGGCACTGGAAGATGATTGACCGAAA 463  
 QY 464 GTGAGGCTTGAAGAGAGATGCTTGTACTTACAAAGCCCTAAGCAGGTGACCA 523  
 DB 464 GTGAGGCTTGAAGAGAGATGCTTGTACTTACAAAGCCCTAAGCAGGTGACCA 523  
 QY 524 GAATGGCTGCCAAGCTGCATCCATGACAAACGCAAAAGTGGCCAGAGCTTGCAAGTT 583  
 DB 524 GAATGGCTGCCAAGCTGCATCCATGACAAACGCAAAAGTGGCCAGAGCTTGCAAGTT 583  
 QY 584 TTGAGAGAACGAGATCTCTCATGATGGAATTCT 618  
 DB 584 TTGAGAGAACGAGATCTCTCATGATGGAATTCT 618

RESULT 6  
 ABL68710/c

ID ABL68710 standard; DNA: 355 BP.

XX ABL68710;

XX 15-MAY-2002 (first entry)

DE Kidney cancer related gene sequence SEQ ID NO:7047.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;

KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;

XX gene; ds.

XX Homo sapiens.

PN WO200194629-A2.

XX 13-DEC-2001.

PF 30-MAY-2001; 2001MO-US10838.

PR 05-JUN-2000; 2000US-209473P.

PR 05-JUN-2000; 2000US-209531P.

PR 18-SEP-2000; 2000US-233133P.

PR 18-SEP-2000; 2000US-233617P.

PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.

PR 20-SEP-2000; 2000US-234052P.

PR 22-SEP-2000; 2000US-234567P.

PR 22-SEP-2000; 2000US-234923P.

PR 25-SEP-2000; 2000US-234924P.

PR 25-SEP-2000; 2000US-235077P.

PR 25-SEP-2000; 2000US-235082P.

PR 25-SEP-2000; 2000US-235134P.

PR 25-SEP-2000; 2000US-235280P.

PR 26-SEP-2000; 2000US-235637P.

PR 26-SEP-2000; 2000US-235638P.

PR 27-SEP-2000; 2000US-235711P.

PR 27-SEP-2000; 2000US-235720P.

PR 27-SEP-2000; 2000US-235840P.

PR 27-SEP-2000; 2000US-235863P.

```

PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 28-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 29-SEP-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI: 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1; SEQ ID 7047; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (1) of a signature gene set, where (1)
XX comprises a sequence (S) selected from 8447 sequences (given in AB61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (1) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which
XX is the data collected with respect to the anti-neoplastic agent as a
XX result of M1, and the data is sufficient to convey the chemical
XX structure and/or properties of the agent. M1 can be used in the
XX treatment of cancer such as colon, breast, stomach, lung, thyroid,
XX oesophageal, ovarian, kidney, prostate or pancreatic cancer;
XX adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
XX infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
XX carcinoma, papillary carcinoma and Wilms' tumour.
XX
XX Sequence 355 BP; 112 A; 74 C; 54 G; 115 T; 0 other:
XX
XX Query Match 11.8%; Score 241; DB 24; Length 355;
XX Best Local Similarity 99.4%; Pred. No. 9.5e-10;
XX Matches 341; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1690 AAAGAGTTTATGTCGCCGACTGCGCAAAATATCATATTCACAGTCTTTGTAG 1749
XX |
XX 343 AAAGAAGTTTATGTCGCCGACTGCGCAAAATATTCATTAATTCACAGTCTTTGTAG 284
XX |
XX 1750 ATGACTGAAGTTTATGTCGACCATATTTGGAGTTTGAATTTAGTGAATGGCAGGAA 1809
XX |
XX 283 ATGACTGAAGTTTATGTCGACCATATTTGGAGTTTGAATTTAGTGAATGGCAGGAA 224
XX |
XX 1810 AGGCGCATCTCATGAGATGATTAAGTGAACCAACTAGTCTCGGATTTTACAGAGA 1869
XX |
XX 223 AGCGCATCTCATGAGATGATTAAGTGAACCAACTAGTCTCGGATTTTACAGAGA 164
XX |
XX 1870 AGGAGGATCTGACTGAGGAGAGTGTGACATGAGACTTGAGACCAAAAGACTTTGAAT 1929
XX |

```

```

DB 163 AGGAGGATCTGACTGAGGAGAGTGTGACATGAGACTTTGAAGACCAAGACTTTGAAT 104
OY 1930 TTGCAGCTGCTCATGCTGAGTTTATATCACTGCTGCTTTCTTATGATTACAAATCT 1989
DB 103 TTGCAGCTGCTCATGCTGAGTTTATATCACTGCTGCTTTCTTATGATTACAAATCT 44
OY 1990 ATATTTTATGCACTTTAAATAAGAAAAAATTTTACAGAAA 2032
DB 43 ATATTTTATGCACTTTAAATAAGAAAAAATTTTACAGAAA 1
XX
XX RESULT 7
XX AAA00911
XX ID AAA00911 standard; cDNA; 300 BP.
XX
XX AAA00911;
XX
XX 19-MAY-2000 (first entry)
XX
XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:902.
XX
XX Human; colon cancer; tumour; diagnosis; gene expression product;
XX probe; detection; cancerous state; metastasis; identification;
XX breast cancer; oestrogen receptor-positive breast cancer; therapy;
XX oestrogen receptor-negative breast cancer; lung cancer; ss.
XX
XX Homo sapiens.
XX
XX WO9958675-A2.
XX
XX 18-NOV-1999.
XX
XX 13-MAY-1999; 99WO-US10602.
XX
XX 14-MAY-1998; 98US-0085426.
XX 15-MAY-1998; 98US-0085537.
XX 15-MAY-1998; 98US-0085696.
XX 21-OCT-1998; 98US-0105234.
XX 27-OCT-1998; 98US-0105877.
XX
XX (CHIR) CHIRON CORP.
XX (HYSE-) HYSEQ INC.
XX
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;
XX Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A;
XX Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;
XX
XX WPI: 2000-126369/11.
XX
XX Polynucleotide library used to determine cancerous states of mammalian
XX cells -
XX
XX Claim 1; Page 399; 1097pp; English.
XX
XX AAA00010 to AAA02716 represent polynucleotides isolated from cDNA
XX libraries constructed from human colon cancer cell lines. The present
XX invention also describes a method of detecting differentially expressed
XX genes correlated with a cancerous state of a mammalian cell, comprising
XX detecting at least one differentially expressed gene product in a test
XX sample derived from a cell suspected of being cancerous, where detection
XX of the differentially expressed gene product is correlated with a
XX cancerous state of the cell from which the test sample was derived.
XX The polynucleotides sequences can be used in a method for detecting
XX differentially expressed genes correlated with a cancerous state of a
XX mammalian cell. The polynucleotides can also be used as probes for
XX detecting and mapping related genes. They can be used in diagnosis and
XX prognosis of diseases and disorders (e.g. identification of
XX pre-metastatic or metastatic cancerous states, stages of cancer, or
XX responsiveness of cancer to therapy). This is particularly for breast
XX cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-
XX negative breast cancer, lung cancer, and colon cancer.
XX

```

SQ Sequence 300 BP: 83 A; 69 C; 72 G; 76 T; 0 other:

Query Match 8.6%; Score 175; DB 21; Length 300;  
Best Local Similarity 99.6%; Pred. No. 1.7e-72;  
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 939 GACCTGGTCCCATGTCCTCCCTGCTATGAGCTAGAGATCATGATGCTCGAAGGCG 998  
|||||  
Db 75 GACCTGGTCCCATGTCCTCCCTGCTATGAGCTAGAGATCATGATGCTCGAAGTGGG 134  
|||||

QY 999 AGGAGTCTGTTCTTGAACCTGCTCTTGAATCGTCAAGTTTCATCCAGGCCACAGC 1058  
|||||  
Db 135 AAGAGTCTGTTCTTGAACCTGCTCTTGAATCGTCAAGTTTCATCCAGGCCACAGC 194  
|||||

QY 1059 CTACAGCCACTCCCAATTAAGTGCATACATGAGCTGAGACACAGAGATTATACC 1118  
|||||  
Db 195 CTACAGCCACTCCCAATTAAGTGCATACATGAGCTGAGACACAGAGATTATACC 254  
|||||

QY 1119 TGTGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCG 1164  
|||||  
Db 255 TGTGTGACCTCTGTGATGCAATCATCATTTGGGATCGCGAATGGCG 300  
|||||

RESULT 8  
ABA42839/c  
ID ABA42839 standard; DNA: 457 BP.  
XX ABA42839;  
XX  
XX 01-FEB-2002 (first entry)  
XX  
XX Human breast cell single exon nucleic acid probe #1534.  
XX  
XX Human: microarray; single exon probe; gene expression; breast;  
XX disease; cancer; ss.  
XX  
XX Homo sapiens.  
XX  
XX W0200157271-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00662.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR:  
XX  
XX WPI; 2001-496933/54.  
XX  
XX  
XX New spatially-addressable set of single exon nucleic acid probes,  
XX useful for measuring gene expression in sample derived from human  
XX breast, comprises number of single exon nucleic acid probes -  
XX  
XX  
XX Claim 1; SEQ ID NO 1534; 327bp + sequence listing; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
XX nucleic acid probes for measuring gene expression in a sample derived  
XX from human breast and BT 474 cells. The method involves contacting  
XX the probes with a collection of detectably labelled nucleic acids  
XX derived from mRNA of human breast, and then measuring the label  
XX bound to each probe of the microarray. The probes are useful for  
XX verifying the expression of regions of genomic DNA predicted to  
XX encode proteins. They are useful for gene discovery, and for  
XX determining predisposition and/or prognosing breast disease. Gene

CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIP0 at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 457 BP: 120 A; 102 C; 110 G; 125 T; 0 other:

Query Match 7.1%; Score 145; DB 22; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.8e-58;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 AGGTCTATGAAGCCTGAGCATCATGACCAACAAGTTTCTGCCCAAGACAGAAATCT 242  
|||||  
Db 450 AGGTCTATGAAGCCTGAGCATCATGACCAACAAGTTTCTGCCCAAGACAGAAATCT 391  
|||||

QY 243 GCCGGCACCAATGATCAGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 302  
|||||  
Db 390 GCCGGCACCAATGATCAGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 331  
|||||

QY 303 TCAGAAATAGACCAACTGCTCTGAT 327  
|||||  
Db 330 TCAGAAATAGACCAACTGCTCTGAT 306  
|||||

RESULT 9  
ABA53262/c  
ID ABA53262 standard; DNA: 457 BP.  
XX ABA53262;  
XX  
XX 01-FEB-2002 (first entry)  
XX  
XX Human foetal liver single exon nucleic acid probe #1567.  
XX  
XX Human: foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX  
XX Homo sapiens.  
XX  
XX W0200157277-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00669.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR:  
XX  
XX WPI; 2001-483447/52.  
XX  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human foetal liver -  
XX  
XX  
XX Claim 1; SEQ ID NO 1567; 639bp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
XX liver. The single exon nucleic acid probes may be used for predicting,  
XX measuring and displaying gene expression in samples derived from human  
XX foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [http://wipo.int/pub/published\\_pcc\\_sequences](http://wipo.int/pub/published_pcc_sequences).  
XX

SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.8e-58;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 AGGCTATGAGGCGCTAGACATCATCACCACAAAGGTTTCTGCCAAGAGAGAGAATCT 242  
|||||  
DB 450 AGGCTATGAGGCGCTAGACATCATCACCACAAAGGTTTCTGCCAAGAGAGAGAATCT 391  
OY 243 GCCGGCACCACATGATGATCAGCTTTGTGATCCTCTTTGTACCAATTACACAGTGGTGA 302  
|||||  
DB 390 GCCGGCACCACATGATGATCAGCTTTGTGATCCTCTTTGTACCAATTACACAGTGGTGA 331  
OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327  
|||||  
DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306

RESULT 10  
ID AAK01519/c  
ID AAK01519 standard; DNA; 457 BP.  
XX  
XX AAK01519;  
AC  
XX  
XX  
DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe SEQ ID NO: 1510.

XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
XX  
OS Homo sapiens.

PN WO200157275-A2.

XX  
XX  
PD 09-AUG-2001.

XX  
XX  
PF 30-JAN-2001; 2001WO-US00667.

XX  
XX  
PR 04-FEB-2000; 2000US-0180312.

XX  
XX  
PR 26-MAY-2000; 2000US-0207456.

XX  
XX  
PR 30-JUN-2000; 2000US-0608408.

XX  
XX  
PR 03-AUG-2000; 2000US-0632366.

XX  
XX  
PR 21-SEP-2000; 2000US-0234687.

XX  
XX  
PR 27-SEP-2000; 2000US-0236359.

XX  
XX  
PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
XX  
DR WPI; 2001-483446/52.

PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -

PS Example 4; SEQ ID NO: 1510; 650PP + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is one of the probes of the  
XX invention.

SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.8e-58;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 AGGCTATGAGGCGCTAGACATCATCACCACAAAGGTTTCTGCCAAGAGAGAGAATCT 242  
|||||  
DB 450 AGGCTATGAGGCGCTAGACATCATCACCACAAAGGTTTCTGCCAAGAGAGAGAATCT 391  
OY 243 GCCGGCACCACATGATGATCAGCTTTGTGATCCTCTTTGTACCAATTACACAGTGGTGA 302  
|||||  
DB 390 GCCGGCACCACATGATGATCAGCTTTGTGATCCTCTTTGTACCAATTACACAGTGGTGA 331  
OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327  
|||||  
DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306

RESULT 11  
ID AAK26966/c  
ID AAK26966 standard; DNA; 457 BP.  
XX  
XX  
XX AAK26966;  
AC  
XX  
XX  
DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 1523.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukemia; lymphoma; myeloma; ss.  
XX  
OS Homo sapiens.

PN WO200157276-A2.

XX  
XX  
PD 09-AUG-2001.

XX  
XX  
PF 30-JAN-2001; 2001WO-US00668.

XX  
XX  
PR 04-FEB-2000; 2000US-0180312.

XX  
XX  
PR 26-MAY-2000; 2000US-0207456.

XX  
XX  
PR 30-JUN-2000; 2000US-0608408.

XX  
XX  
PR 03-AUG-2000; 2000US-0632366.

XX  
XX  
PR 21-SEP-2000; 2000US-0234687.

XX  
XX  
PR 27-SEP-2000; 2000US-0236359.

XX  
XX  
PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
XX  
DR WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 1523; 658PP + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukemia and myeloma. The present sequence is one of  
XX the probes of the invention.

SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.8e-58;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 183 AGGCTATGAGGCGCTAGACATCATCACCACAAAGGTTTCTGCCAAGAGAGAGAATCT 242

|||||  
Db 450 AGGCTATGAGGCGCTAGCATCATCCAAAGAGTTTGTGCCACAGACAGAGATCT 391  
Oy 243 GCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 302  
Db 390 GCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 331  
Oy 303 TCAGAAATAGAGCACTGCTCTGAT 327  
Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

RESULT 12  
AA11574/c  
ID AA11574 standard; DNA: 457 BP.  
XX  
AC AA11574;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Probe #1507 for gene expression analysis in human cervical cell sample.  
XX  
KW Probe: human; microarray; gene expression; cervical epithelial cell;  
XX  
OS cervical cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00670.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000CB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI: 2001-488901/53.  
XX  
XX

Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human cervical epithelial cells -  
Claim 25; SEQ ID No 1507; 487pp; English.

The present invention relates to human single exon nucleic acid probes  
(SENP). The present sequence is one such probe. The SENPs are derived  
from human HeLa cells. The SENPs can be used to produce a single exon  
microarray, which can be used for measuring human gene expression in a  
sample derived from human cervical epithelial cells. By measuring gene  
expression, the probes are therefore useful in grading and/or staging  
of diseases of the cervix, notably cervical cancer.  
Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.8e-58;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 183 AGGCTATGAGGCGCTAGCATCATCCAAAGAGTTTGTGCCACAGACAGAGATCT 242  
Db 450 AGGCTATGAGGCGCTAGCATCATCCAAAGAGTTTGTGCCACAGACAGAGATCT 391

Oy 243 GCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 302  
Db 390 GCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 331  
Oy 303 TCAGAAATAGAGCACTGCTCTGAT 327  
Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

RESULT 13  
AA132862/c  
ID AA132862 standard; DNA: 457 BP.  
XX  
AC AA132862;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Probe #1548 used to measure gene expression in human placenta sample.  
XX  
KW Probe: microarray; human; placenta; antenatal diagnosis;  
XX  
OS genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
XX  
PR 26-MAY-2000; 2000US-0207456.  
XX  
PR 30-JUN-2000; 2000US-0608408.  
XX  
PR 03-AUG-2000; 2000US-0632366.  
XX  
PR 21-SEP-2000; 2000US-0234687.  
XX  
PR 27-SEP-2000; 2000US-0236359.  
XX  
PR 04-OCT-2000; 2000CB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI: 2001-488897/53.  
XX  
XX

Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human placenta -  
Claim 25; SEQ ID No 1548; 654pp; English.

The present invention relates to single exon nucleic acid probes (SENP).  
The present sequence is one such probe. The probes are useful for  
producing a microarray for predicting, measuring and displaying gene  
expression in samples derived from human placenta. The probes are useful  
for antenatal diagnosis of human genetic disorders.

Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match 7.1%; Score 145; DB 22; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2.8e-58;  
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 183 AGGCTATGAGGCGCTAGCATCATCCAAAGAGTTTGTGCCACAGACAGAGATCT 242  
Db 450 AGGCTATGAGGCGCTAGCATCATCCAAAGAGTTTGTGCCACAGACAGAGATCT 391  
Oy 243 GCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 302  
Db 390 GCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGACT 331  
Oy 303 TCAGAAATAGAGCACTGCTCTGAT 327  
Db 330 TCAGAAATAGAGCACTGCTCTGAT 306

## RESULT 14

ABA47966/C

ID ABA47966 standard; DNA: 138 BP.

AC ABA47966;

DT 01-FEB-2002 (first entry)

XX Human breast cell single exon nucleic acid probe #6661.

XX Human: microarray: single exon probe; gene expression: breast;

XX disease: cancer; ss.

XX Homo sapiens.

OS WO200157271-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00662.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

XX WPI: 2001-496933/54.

DB 78 CGGCACACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTAACAGAGTGAGACTTC 19  
OY 305 AGAATATAGACCACTGCT 322  
DB 18 AGAATATAGACCACTGCT 1

## RESULT 15

ABA65854/C

ID ABA65854 standard; DNA: 138 BP.

AC ABA65854;

DT 01-FEB-2002 (first entry)

XX Human foetal liver single exon nucleic acid probe #14159.

XX Human: foetal liver; gene expression: single exon nucleic acid probe; ss.

XX Homo sapiens.

OS WO200157271-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

XX WPI: 2001-483447/52.

SO Sequence 138 BP; 32 A; 31 C; 35 G; 40 T; 0 other;  
Query Match 6.8%; Score 138; DB 22; Length 138;  
Best Local Similarity 100.0%; Pred. No. 5.8e-55;  
Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 185 GTCTATGAAGCCTTAGACATCATCACCACAAAGTTTCTGCCCAAGAGCAGAAATCTGC 244  
DB 138 GTCTATGAAGCCTTAGACATCATCACCACAAAGTTTCTGCCCAAGAGCAGAAATCTGC 79  
OY 245 CGGCACACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTAACAGAGTGAGACTTC 304  
DB 78 CGGCACACATGATCAGCTTTGTGGATCCTCTTGTGACCAATTAACAGAGTGAGACTTC 19  
OY 305 AGAATATAGACCACTGCT 322  
DB 18 AGAATATAGACCACTGCT 1



Tue Apr 22 09:11:13 2003

us-09-513-151-3.oli.rng

Page 14

Search completed: April 21, 2003, 21:21:56  
Job time : 485 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 19:01:40 ; Search time 20 Seconds

(without alignments)  
6005.218 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCCATACATGCGCTCCG.....TTTACACAGAAAAA 2041

Scoring table:

	OLIGO	Xgapop	Xgapext	Ygapop	Ygapext	Fgapop	Fgapext	Delop
		60.0	60.0	60.0	60.0	6.0	7.0	6.0
						6.0	7.0	7.0

Searched: 262574 seqs, 29422922 residues

Word size: 1

Total number of hits satisfying chosen parameters: 438908

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE1=frame\_n2p\_model -DEV=xlh  
-O=/cgn2\_1/USPO.spool/US09513151/runat.15042003.141247.27244/app\_query.fasta.1.2183  
-DB-Issued\_Patents\_AA -QFMT=fastan -SUFFIX=oln2p.ra1 -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=DITS -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09513151 -CGN1.1.16 -runat.15042003.141247.27244 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -TRENDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELEX=7

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	8	1.2	378	4	US-09-215-252-17
2	8	1.2	505	4	US-09-509-902A-11
3	8	1.2	666	4	US-09-134-001C-5465
4	8	1.2	1087	1	US-08-264-002-5
5	8	1.2	1306	4	US-08-999-774A-13
6	8	1.1	15	2	US-08-686-594-14
7	8	1.1	17	2	US-08-686-594-4
8	8	1.1	18	2	US-08-686-594-1
9	8	1.1	18	2	US-08-686-594-5
10	8	1.1	18	2	US-08-686-594-11
11	8	1.1	19	2	US-08-686-594-2
12	8	1.1	19	2	US-08-686-594-6

C 13	7	1.1	19	2	US-08-686-594-12
C 14	7	1.1	20	2	US-08-686-594-3
C 15	7	1.1	20	2	US-08-686-594-7
C 16	7	1.1	20	2	US-08-686-594-13
C 17	7	1.1	26	1	US-07-942-245-287
C 18	7	1.1	26	1	US-07-942-245-325
C 19	7	1.1	26	1	US-07-942-245-348
C 20	7	1.1	73	4	US-09-134-001C-5563
C 21	7	1.1	83	4	US-09-100-802-6
C 22	7	1.1	104	1	US-08-920-440B-8
C 23	7	1.1	104	2	US-08-920-440B-8
C 24	7	1.1	104	4	US-09-173-492-8
C 25	7	1.1	104	4	US-09-173-132-8
C 26	7	1.1	104	4	US-09-165-533-8
C 32	7	1.1	116	1	US-08-209-747-18
C 33	7	1.1	116	1	US-08-458-298-18
C 34	7	1.1	129	6	5428135-6
C 35	7	1.1	132	4	US-09-034-088A-4
C 36	7	1.1	133	4	US-09-452-229-10
C 37	7	1.1	151	2	US-08-913-477-11
C 38	7	1.1	169	4	US-09-134-001C-4614
C 39	7	1.1	177	4	US-09-643-597-165
C 40	7	1.1	184	1	US-09-609-324A-2
C 41	7	1.1	184	2	US-08-920-440B-2
C 42	7	1.1	184	4	US-09-173-492-2
C 43	7	1.1	184	4	US-09-173-132-2
C 44	7	1.1	184	4	US-09-165-533-2
C 45	7	1.1	184	5	PCR-US95-12779-2

#### ALIGNMENTS

RESULT 1  
US-09-215-252-17  
Sequence 17, Application US/09215252  
Patent No. 6300487  
GENERAL INFORMATION:  
APPLICANT: LEUNG, David W.  
APPLICANT: ADONEL, Daniel  
TITLE OF INVENTION: MAMMALIAN LYSOPHOSPHATIDIC ACID ACYLTRANSFERASE  
FILE REFERENCE: 077319/0151  
CURRENT APPLICATION NUMBER: US/09/215,252  
PRIOR FILING DATE: 1998-12-18  
PRIOR APPLICATION NUMBER: US 08/618,651  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 17  
LENGTH: 378  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-215-252-17

Alignment Scores:  
Pred. No.: 21.9  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.21%  
DB: 4  
Gaps: 0

US-09-513-151-3 (1-2041) x US-09-215-252-17 (1-378)  
QY 669 AGTCTCTAACCCTGATCCTT 692  
DB 334 SerSerLeuThrLeuAlaSerPhe 341

Sequence 12, Appl  
Sequence 3, Appl  
Sequence 7, Appl  
Sequence 13, Appl  
Sequence 287, App  
Sequence 325, App  
Sequence 348, App  
Sequence 5563, Ap  
Sequence 6, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 29, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 18, Appl  
Patent No. 5428135  
Sequence 4, Appl  
Sequence 10, Appl  
Sequence 11, Appl  
Sequence 165, App  
Sequence 166, App  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl  
Sequence 2, Appl

```
RESULT 2
US-09-509-902A-11
; Sequence 11, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Maiken, John S.
; TITLE OF INVENTION: Human CDNs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-509-902A-11

Alignment Scores:
Pred. No.: 21          Length: 505
Score: 8.00           Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21%     Indels: 0
DB: 4                  Gaps: 0

US-09-513-151-3 (1-2041) x US-09-509-902A-11 (1-505)
OY 17 TCCGGGGCGGCTGCAGCAGT 40
DB 496 Serva1A1A1A1A1A1A1A1 503

RESULT 3
US-09-134-001C-5465
; Sequence 5465, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5465
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5465

Alignment Scores:
Pred. No.: 20.2        Length: 666
Score: 8.00           Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21%     Indels: 0
DB: 4                  Gaps: 0

US-09-513-151-3 (1-2041) x US-09-134-001C-5465 (1-666)
OY 95 CTCGGGGCCAGCGGCAGCGCAA 118
DB 45 LeuG1A1A1A1A1A1A1A1A1 52

RESULT 4
US-08-264-002-5
```

```
; Sequence 5, Application US/08264002
; Patent No. 5559019
; GENERAL INFORMATION:
; APPLICANT: GUI, JIAN-FANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1087 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: CEHK
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..1087
US-08-264-002-5

Alignment Scores:
Pred. No.: 18.8        Length: 1087
Score: 8.00           Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.22%     Indels: 0
DB: 1                  Gaps: 0

US-09-513-151-3 (1-2041) x US-08-264-002-5 (1-1087)
OY 1910 TCAAGCTATGTCACAGCTTCT 1887
DB 829 SerSerPrometSerGlnLeupro 836

RESULT 5
US-08-999-774A-13
; Sequence 13, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
```

STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/999,774A  
FILING DATE: 10-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US\*60/032,818  
FILING DATE: 11-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0646  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)496-1200  
TELEFAX: (650)852-9196  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1306 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-999-774A-13

Alignment Scores:  
Pred. No.: 18.4 Length: 1306  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.22% Indels: 0  
DB: 4 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-999-774A-13 (1-1306)

QY 1307 TCTGCCCTGGGATCCCTCCCTT 1284  
|||||  
DB 637 SerialaLeuGlyIleProSerLeu 644

RESULT 6  
US-08-686-594-14  
Sequence 14, Application US/08686594  
Patent No. 5856127  
GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.  
APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:

NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-14

Alignment Scores:  
Pred. No.: 33.2 Length: 15  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.06% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-14 (1-15)

QY 152 CGAGCCGCTGGCTAGCTGCA 132  
|||||  
DB 2 ArgAlaAlaGlyLeuAlaAla 8

RESULT 7  
US-08-686-594-4  
Sequence 4, Application US/08686594  
Patent No. 5856127  
GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.  
APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-4

Alignment Scores:  
Pred. No.: 32.6 Length: 17  
Score: 7.00 Matches: 7

Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.06%  
DB: 2  
Conservative: 0  
Matches: 0  
Indels: 0  
Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-4 (1-17)

QY 152 CGAGCCGCTGGCTAGCTGCA 132

DB 3 ArgAlaAlaGlyLeuAlaAla 9

RESULT 8

US-08-686-594-1  
Sequence 1, Application US/08686594  
Patent No. 5856127

GENERAL INFORMATION:

APPLICANT: POWELL, WILLIAM A.

TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP

STREET: P.O. BOX 1051, CLINTON SQUARE

CITY: ROCHESTER

STATE: NEW YORK

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/686,594

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: TIMIAN, SUSAN J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 20884/50

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1636

TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-686-594-1

#### Alignment Scores:

Pred. No.: 323  
Score: 7.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.06%  
Length: 18  
Matches: 7  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-1 (1-18)

QY 152 CGAGCCGCTGGCTAGCTGCA 132

DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 9

US-08-686-594-5  
Sequence 5, Application US/08686594  
Patent No. 5856127

GENERAL INFORMATION:

APPLICANT: POWELL, WILLIAM A.

APPLICANT: MAYNARD, CHARLES A.

TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP

STREET: P.O. BOX 1051, CLINTON SQUARE

CITY: ROCHESTER

STATE: NEW YORK

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/686,594

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: TIMIAN, SUSAN J.

REGISTRATION NUMBER: 34,103

REFERENCE/DOCKET NUMBER: 20884/50

TELECOMMUNICATION INFORMATION:

TELEPHONE: 716-263-1636

TELEFAX: 716-263-1600

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-686-594-5

#### Alignment Scores:

Pred. No.: 323  
Score: 7.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 1.06%  
Length: 18  
Matches: 7  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-5 (1-18)

QY 152 CGAGCCGCTGGCTAGCTGCA 132

DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 10

US-08-686-594-11  
Sequence 11, Application US/08686594  
Patent No. 5856127

GENERAL INFORMATION:

APPLICANT: POWELL, WILLIAM A.

TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP

STREET: P.O. BOX 1051, CLINTON SQUARE

CITY: ROCHESTER

STATE: NEW YORK

COUNTRY: USA

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/686,594

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1636  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-11

Alignment Scores:  
Pred. No.: 323 Length: 18  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.06% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-11 (1-18)

QY 152 CGAGCCGCTGGCTAGCTGCA 132  
DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 11  
US-08-686-594-2  
Sequence 2, Application US/08686594  
Patent No. 5856127

GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.  
APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1600  
TELEFAX: 716-263-1636  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-2

Alignment Scores:  
Pred. No.: 321 Length: 19

Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.06% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-2 (1-19)

QY 152 CGAGCCGCTGGCTAGCTGCA 132  
DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 12  
US-08-686-594-6  
Sequence 6, Application US/08686594  
Patent No. 5856127

GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.  
APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1600  
TELEFAX: 716-263-1636  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-6

Alignment Scores:  
Pred. No.: 321 Length: 19  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.06% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-6 (1-19)

QY 152 CGAGCCGCTGGCTAGCTGCA 132  
DB 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 13  
US-08-686-594-12  
Sequence 12, Application US/08686594  
Patent No. 5856127

GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.

APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: \*  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1600  
TELEFAX: 716-263-1636  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-12

Alignment Scores:  
Pred. No.: 321 Length: 19  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.06% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-12 (1-19)  
OY 152 CGAGCCGCTGGCTAGTCGA 132  
Db 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 14  
US-08-686-594-3  
Sequence 3, Application US/08686594  
Patent No. 5856127  
GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.  
APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1600  
TELEFAX: 716-263-1636  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-3

Alignment Scores:  
Pred. No.: 319 Length: 20  
Score: 7.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.06% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x US-08-686-594-3 (1-20)  
OY 152 CGAGCCGCTGGCTAGTCGA 132  
Db 4 ArgAlaAlaGlyLeuAlaAla 10

RESULT 15  
US-08-686-594-7  
Sequence 7, Application US/08686594  
Patent No. 5856127  
GENERAL INFORMATION:  
APPLICANT: POWELL, WILLIAM A.  
APPLICANT: MAYNARD, CHARLES A.  
TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON, HARGRAVE, DEVANS & DOYLE LLP  
STREET: P.O. BOX 1051, CLINTON SQUARE  
CITY: ROCHESTER  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/686,594  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: TIMIAN, SUSAN J.  
REGISTRATION NUMBER: 34,103  
REFERENCE/DOCKET NUMBER: 20884/50  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1600  
TELEFAX: 716-263-1636  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-686-594-7

Alignment Scores:

Pred. No.:	319	Length:	20
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.06%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x US-08-686-594-7 (1-20)

QY 152 CGAGCCGCTGGCCTAGCTGCA 132  
|||||

Db 4 ArgAlaAlaGlyLeuAlaAla 10

Search completed: April 21, 2003, 19:19:23  
Job time : 26 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 16:30:20 ; Search time 71.8325 seconds

(without alignments)  
7572.189 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575

Sequence: 1 CTGCCATAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Dgapop 6.0 , Dgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODFL=frame+np2 model -DEV=ylh  
-Q=/cgn2.1/USPTO.spool/US09513151/Runat.15042003\_141143\_26358/app.query.fasta\_1.2446  
-DB-A.Geneseq\_101002 -OFMT=fasta -SOFITX=n2p.rag -MINMATCH=0.1 -LOOPTL=0  
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151\_@cgn2.1\_1.84\_@runat.15042003\_141143\_26358 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONCLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq\_101002:\*

- 1: /SID52/gcgdata/geneseq/geneseqp.embl/AA1980.DAT:\*
- 2: /SID52/gcgdata/geneseq/geneseqp.embl/AA1981.DAT:\*
- 3: /SID52/gcgdata/geneseq/geneseqp.embl/AA1982.DAT:\*
- 4: /SID52/gcgdata/geneseq/geneseqp.embl/AA1983.DAT:\*
- 5: /SID52/gcgdata/geneseq/geneseqp.embl/AA1984.DAT:\*
- 6: /SID52/gcgdata/geneseq/geneseqp.embl/AA1985.DAT:\*
- 7: /SID52/gcgdata/geneseq/geneseqp.embl/AA1986.DAT:\*
- 8: /SID52/gcgdata/geneseq/geneseqp.embl/AA1987.DAT:\*
- 9: /SID52/gcgdata/geneseq/geneseqp.embl/AA1988.DAT:\*
- 10: /SID52/gcgdata/geneseq/geneseqp.embl/AA1989.DAT:\*
- 11: /SID52/gcgdata/geneseq/geneseqp.embl/AA1990.DAT:\*
- 12: /SID52/gcgdata/geneseq/geneseqp.embl/AA1991.DAT:\*
- 13: /SID52/gcgdata/geneseq/geneseqp.embl/AA1992.DAT:\*
- 14: /SID52/gcgdata/geneseq/geneseqp.embl/AA1993.DAT:\*
- 15: /SID52/gcgdata/geneseq/geneseqp.embl/AA1994.DAT:\*
- 16: /SID52/gcgdata/geneseq/geneseqp.embl/AA1995.DAT:\*
- 17: /SID52/gcgdata/geneseq/geneseqp.embl/AA1996.DAT:\*
- 18: /SID52/gcgdata/geneseq/geneseqp.embl/AA1997.DAT:\*
- 19: /SID52/gcgdata/geneseq/geneseqp.embl/AA1998.DAT:\*
- 20: /SID52/gcgdata/geneseq/geneseqp.embl/AA1999.DAT:\*
- 21: /SID52/gcgdata/geneseq/geneseqp.embl/AA2000.DAT:\*
- 22: /SID52/gcgdata/geneseq/geneseqp.embl/AA2001.DAT:\*
- 23: /SID52/gcgdata/geneseq/geneseqp.embl/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1989	55.6	411	23	AB97295
2	1074	30.0	222	22	AB10278
3	1074	30.0	222	22	AB10476
4	1074	30.0	222	22	AA23398
5	549	15.4	430	20	AA02526
6	435	12.2	221	20	AA102534
7	397.5	11.1	305	23	AB48079
8	355	9.9	329	22	AA682183
9	338	9.5	318	23	AB93084
10	334.5	9.4	311	22	AA88507
11	334.5	9.4	311	22	AA91838
12	333.5	9.3	329	21	AA49765
13	333.5	9.3	329	23	AB892405
14	333	9.3	294	23	AB85336
15	327.5	9.2	290	21	AA649766
16	327.5	9.2	330	23	AB93538
17	326	9.1	330	23	AB92364
18	322.5	9.0	357	23	AB91502
19	319.5	8.9	329	21	AA61729
20	319.5	8.9	342	20	AA35565
21	318.5	8.9	299	23	AB26718
22	313.5	8.8	290	21	AA617330
23	313.5	8.8	330	23	AB92775
24	307.5	8.6	193	23	AB39326
25	307	8.6	300	23	AB26717
26	301.5	8.4	301	22	AA691879
27	301.5	8.4	463	21	AA620250
28	299	8.4	463	21	AB93051
29	298	8.3	447	23	AB93360
30	297	8.3	447	23	AB93360
31	288.5	8.1	357	22	AA057193
32	267	7.5	342	23	AB91127
33	246	6.9	57	22	AB828679
34	246	6.9	57	22	AB834055
35	246	6.9	57	22	AA54837
36	246	6.9	57	22	AA667220
37	246	6.9	57	22	AA615066
38	246	6.9	57	22	AA627510
39	243	6.8	396	21	AA620251
40	243	6.8	396	21	AA622765
41	238.5	6.7	242	21	AB40607
42	238.5	6.7	242	23	AB31897
43	167	4.7	218	19	AA698315
44	132	3.7	87	20	AAV37377
45	125	3.5	220	20	AAV37376

# ALIGNMENTS

RESULT 1

ID ABB97295 standard; Protein; 411 AA.

AC ABB97295;

DT 27-JUN-2002 (first entry)

XX

DE Novel human protein SEQ ID NO: 563.

XX

XX Human; antineoplastic; vulnary; antiinflammatory; immunomodulator;

XX antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;

KW neuroprotective; antiparkinsonian; protein therapy; EST;

XX expressed sequence tag.

XX

OS Homo sapiens.

XX

XX WO200222660-A2.

XX

PD 21-MAR-2002.  
 XX 10-SEP-2001: 2001WO-US26015.  
 PF 11-SEP-2000: 2000US-0659671.  
 PR (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 P1 Xue AD, Yang Y, Wehrman T, Drmanac RT;  
 XX WPI: 2002-292408/33.  
 DR N-PSDB: ABN32481.  
 XX An isolated polynucleotide for treating diseases associated with its  
 FT encoded polypeptide such as cancer and multiple sclerosis.  
 PT  
 XX Example 2: SEQ ID NO 563: 509pp; English.  
 XX The present invention provides the protein and coding sequences of 444  
 CC novel human proteins. These were isolated from expressed sequences tags  
 CC (ESTs). They can be used to stimulate cell growth, to regulate  
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
 CC multiple sclerosis, to regulate actin or inhibit e.g. to treat  
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
 CC Parkinson's disease. The present sequence is a protein of the invention.  
 XX  
 SO Sequence 411 AA:  
 Alignment Scores:  
 Pred. No.: 1,42e-206 Length: 411  
 Score: 1989.00 Matches: 383  
 Percent Similarity: 93.43% Conservative: 1  
 Best Local Similarity: 93.19% Mismatches: 1  
 Query Match: 55.64% Indels: 26  
 DB: Gaps: 1  
 US-09-513-151-3 (1-2041) x ABB97295 (1-411)  
 OY 179 ATGCAGTCTATGAGGCTTAGACATCATCAACCAAGGTTTCCCAAGACGAGAGA 238  
 DB 1 MetGlnValIYrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnIAsnArg 20  
 OY 239 ATCTCCGCGACACATGATCAGCTTTGTGATCCTCTTGTGACCAATTAACACTGGTG 298  
 DB 21 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnIYrThrValVal 40  
 OY 299 GACTTCACAAATAGACCAACTGCTGTGATGACATATATTTGCCCGAGACAAATTCCT 358  
 DB 41 AspPheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIlePro 60  
 OY 359 ATTGTGTGGGAGAACCAATTATTAATCTGCAATCTGCTGGAAGTCTTGTCAAT 418  
 DB 61 IleValValIglYglYThrAsnIYrIleGlnSerLeuLeuIYrPlyValLeuValAsn 80  
 OY 419 ACCAAGCCCCAGAGATGGGCACTGAGAAAGTATGACCAAAAGTGGACTTGAAG 478  
 DB 81 ThrIYrProGlnGlnIleMetGlyThrCduLysValIleAspArgLysValGlnLeuGlnLys 100  
 OY 479 GAGATGGTCTTGTACTTCACAAACGCTTAAGCCAGGTGACCCGAAATGGCTGCCAAG 538  
 DB 101 GluAspArgLysLeuValLeuHisLysArgLeuSerGlnValAspProGlnIleAlaLys 120  
 OY 539 CTGCATCCACATGACAAACGCAAAAGTGGCCAGAGCTTGCAAGTTTGAAGAACAAGCA 598  
 DB 121 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnIYrThrGly 140  
 OY 599 ATCTCTCATAGATTTCTCATGTCATACATAGCAAGGAAGGGGTGGTCCCTTGA 658  
 DB 141 IleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnIYrGlyGlyProLeuGly 160

OY 659 GGTCCCTTGAAGTTCTCTAACCCCTTGATCCTTTGGCTTCATGCTGACACGAGCTTCTA 718  
 DB 161 GlyProLeuLysPheSerAsnProCysIleLeuIYrPheHisAlaAspGlnAlaValLeu 180  
 OY 719 GATGAGCGCTTGATTAAGAGGCTGGATGACATGCTGCTGCTGGGCTTGGAGCAACTA 778  
 DB 181 AspGlnArgLeuAspLysArgValAlaAspSerPheLeuAlaIleGlyLeuLeuGlnIleu 200  
 OY 779 AGAGATTTTTCACAGACGCTATATATCAGAAAGTCTTCGGAATTAATACCAAGACTATCAA 838  
 DB 201 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGln 220  
 OY 839 CATGCTATCTTCCAAATCAATGCGCTTCACAGAAATTTACAGAGTACCTGATCAGAGGA 898  
 DB 221 HisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlnIYrLeuIleThrGlnGly 240  
 OY 899 AAATGCACACGAGAGACTAGTAAACAGCTTAAAGAAGA----- 940  
 DB 241 LysCysThrLeuGlnIYrThrSerAsnGlnLeuLeuLysGlyIleGlnAlaLeuLysGln 260  
 OY 940 ----- 940  
 DB 261 ValThrLysArgTyrAlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArg 280  
 OY 941 CCTGGTCCCATTTGCCCCCTGTATGCTTAAGAGTATGTATGTCGAACTGGAG 1000  
 DB 281 ProGlyProIleValProProValIYrCduLysValSerAspValSerLysIYrGln 300  
 OY 1001 GAGTCTGTTCTTGAACGCGCTTGAATCGTGAAGTTTCATCCAGGCGCACAGCGT 1060  
 DB 301 GluSerValLeuGlnIYrProAlaLeuGlnIleValGlnSerPheIleGlnGlyHisLysPro 320  
 OY 1061 AACGCCACTTCAGACCACTGAGAAAGAAAGAGAGATTGACATGCTGCTCAACACC 1240  
 DB 321 ThrAlaThrProIleLysMetProTyrAsnGlnIleAsnLysArgSerTyrHisIleu 340  
 OY 1121 TGTGACCTCTGTGATCGAATCATCATGTTGGGATGCGCAATGGCGACGGCAATAAATCC 1180  
 DB 341 CysAspLeuCysAspArgIleIleIleGlyAspArgGlnIYrPheAlaHisIleLysSer 360  
 OY 1181 AAATCCCACTTCAGACCACTGAGAAAGAAAGAGAGATTGACATGCTGCTCAACACC 1240  
 DB 361 LysSerHisLysLeuAsnGlnLeuLysLysArgArgArgLeuAspSerAspAlaValAsnThr 380  
 OY 1241 ATGAAAGTCAAGAGTGTTCCTCCAGACTATATACAAAGAACTAAAGGAAGGATCCCA 1300  
 DB 381 IleGlnSerGlnSerValSerProAspHisAsnLysGlnIYrGlySerPro 400  
 OY 1301 GCGCAGATGATCAAGAGCTGAATGCAGCGTT 1333  
 DB 401 GlnGlnAsnAspGlnIleuLysCysSerVal 411  
 RESULT 2  
 ABB10278  
 ID ABB10278 standard; Protein: 222 AA.  
 AC ABB10278:  
 XX 10-JAN-2002 (first entry)  
 XX Human cDNA SEQ ID NO: 586.  
 DE Human  
 XX Human: gene therapy; neural disorder; immune system disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW proliferative disorder; inflammation.  
 OS Homo sapiens.  
 XX  
 XX WO200154474-A2.  
 PN 02-AUG-2001.  
 PD

XX 17-JAN-2001; 2001WO-US01349.  
PF XX  
PR 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 24-FEB-2000; 2000US-184664P.  
PR 02-MAR-2000; 2000US-186350P.  
PR 16-MAR-2000; 2000US-189874P.  
PR 17-MAR-2000; 2000US-190076P.  
PR 18-APR-2000; 2000US-198123P.  
PR 19-MAY-2000; 2000US-205515P.  
PR 07-JUN-2000; 2000US-209467P.  
PR 28-JUN-2000; 2000US-214886P.  
PR 30-JUN-2000; 2000US-215135P.  
PR 07-JUL-2000; 2000US-216647P.  
PR 07-JUL-2000; 2000US-216887P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225213P.  
PR 14-AUG-2000; 2000US-225214P.  
PR 14-AUG-2000; 2000US-225266P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225477P.  
PR 14-AUG-2000; 2000US-225757P.  
PR 14-AUG-2000; 2000US-225758P.  
PR 14-AUG-2000; 2000US-225759P.  
PR 18-AUG-2000; 2000US-226279P.  
PR 22-AUG-2000; 2000US-226681P.  
PR 22-AUG-2000; 2000US-226688P.  
PR 22-AUG-2000; 2000US-227182P.  
PR 23-AUG-2000; 2000US-227009P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-229287P.  
PR 01-SEP-2000; 2000US-229343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 06-SEP-2000; 2000US-230437P.  
PR 06-SEP-2000; 2000US-230438P.  
PR 08-SEP-2000; 2000US-231242P.  
PR 08-SEP-2000; 2000US-231243P.  
PR 08-SEP-2000; 2000US-231244P.  
PR 08-SEP-2000; 2000US-231413P.  
PR 08-SEP-2000; 2000US-231414P.  
PR 08-SEP-2000; 2000US-232080P.  
PR 08-SEP-2000; 2000US-232081P.  
PR 12-SEP-2000; 2000US-231968P.  
PR 14-SEP-2000; 2000US-232397P.  
PR 14-SEP-2000; 2000US-232398P.  
PR 14-SEP-2000; 2000US-232399P.  
PR 14-SEP-2000; 2000US-232401P.  
PR 14-SEP-2000; 2000US-232401P.  
PR 14-SEP-2000; 2000US-233063P.  
PR 14-SEP-2000; 2000US-233064P.  
PR 14-SEP-2000; 2000US-233065P.  
PR 21-SEP-2000; 2000US-234223P.  
PR 21-SEP-2000; 2000US-234274P.  
PR 25-SEP-2000; 2000US-234997P.  
PR 25-SEP-2000; 2000US-234998P.  
PR 26-SEP-2000; 2000US-235484P.  
PR 27-SEP-2000; 2000US-235834P.  
PR 27-SEP-2000; 2000US-235836P.  
PR 29-SEP-2000; 2000US-236327P.  
PR 29-SEP-2000; 2000US-236367P.  
PR 29-SEP-2000; 2000US-236368P.  
  
PR 29-SEP-2000; 2000US-236369P.  
PR 29-SEP-2000; 2000US-236370P.  
PR 02-OCT-2000; 2000US-236802P.  
PR 02-OCT-2000; 2000US-237037P.  
PR 02-OCT-2000; 2000US-237038P.  
PR 02-OCT-2000; 2000US-237039P.  
PR 02-OCT-2000; 2000US-237040P.  
PR 13-OCT-2000; 2000US-239335P.  
PR 13-OCT-2000; 2000US-239337P.  
PR 20-OCT-2000; 2000US-240960P.  
PR 20-OCT-2000; 2000US-241221P.  
PR 20-OCT-2000; 2000US-241785P.  
PR 20-OCT-2000; 2000US-241786P.  
PR 20-OCT-2000; 2000US-241787P.  
PR 20-OCT-2000; 2000US-241808P.  
PR 20-OCT-2000; 2000US-241809P.  
PR 20-OCT-2000; 2000US-241826P.  
PR 01-NOV-2000; 2000US-244617P.  
PR 08-NOV-2000; 2000US-246474P.  
PR 08-NOV-2000; 2000US-246475P.  
PR 08-NOV-2000; 2000US-246476P.  
PR 08-NOV-2000; 2000US-246477P.  
PR 08-NOV-2000; 2000US-246478P.  
PR 08-NOV-2000; 2000US-246523P.  
PR 08-NOV-2000; 2000US-246524P.  
PR 08-NOV-2000; 2000US-246525P.  
PR 08-NOV-2000; 2000US-246526P.  
PR 08-NOV-2000; 2000US-246527P.  
PR 08-NOV-2000; 2000US-246528P.  
PR 08-NOV-2000; 2000US-246532P.  
PR 08-NOV-2000; 2000US-246609P.  
PR 08-NOV-2000; 2000US-246610P.  
PR 08-NOV-2000; 2000US-246611P.  
PR 08-NOV-2000; 2000US-246613P.  
PR 17-NOV-2000; 2000US-249207P.  
PR 17-NOV-2000; 2000US-249208P.  
PR 17-NOV-2000; 2000US-249209P.  
PR 17-NOV-2000; 2000US-249210P.  
PR 17-NOV-2000; 2000US-249211P.  
PR 17-NOV-2000; 2000US-249212P.  
PR 17-NOV-2000; 2000US-249213P.  
PR 17-NOV-2000; 2000US-249214P.  
PR 17-NOV-2000; 2000US-249215P.  
PR 17-NOV-2000; 2000US-249216P.  
PR 17-NOV-2000; 2000US-249217P.  
PR 17-NOV-2000; 2000US-249218P.  
PR 17-NOV-2000; 2000US-249244P.  
PR 17-NOV-2000; 2000US-249245P.  
PR 17-NOV-2000; 2000US-249264P.  
PR 17-NOV-2000; 2000US-249265P.  
PR 17-NOV-2000; 2000US-249297P.  
PR 17-NOV-2000; 2000US-249299P.  
PR 17-NOV-2000; 2000US-249300P.  
PR 01-DEC-2000; 2000US-250160P.  
PR 01-DEC-2000; 2000US-250391P.  
PR 05-DEC-2000; 2000US-251030P.  
PR 05-DEC-2000; 2000US-251988P.  
PR 05-DEC-2000; 2000US-256719P.  
PR 06-DEC-2000; 2000US-251479P.  
PR 08-DEC-2000; 2000US-251856P.  
PR 08-DEC-2000; 2000US-251868P.  
PR 08-DEC-2000; 2000US-251869P.  
PR 08-DEC-2000; 2000US-251989P.  
PR 08-DEC-2000; 2000US-251990P.  
PR 11-DEC-2000; 2000US-254097P.  
PR 05-JAN-2001; 2001US-259678P.  
  
(HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-476161/51.  
XX N-PSDB; ABA06500.  
DR

XX Isolated nucleic acid molecule encoding an inflammation-associated  
PT polypeptide is used in preventing, treating or ameliorating a medical  
PT condition -

PS Claim 11; SEQ ID NO: 586; 859pp + Sequence Listing; English.

CC is a protein of the invention.  
CC renal and proliferative disorders and inflammation. The present sequence.  
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular.  
CC DNAs. These can be used in the treatment of neural, immune system.  
CC The present invention provides human cDNAs, proteins and related genomic  
CC

**SQ**      **Sequence**      **222 AA;**

Alignment Scores:

Pred. No.:	2.68e-107	Length:	222
Score:	1074.00	Matches:	214
Percent Similarity:	98.62%	Conservative:	1
Best Local Similarity:	98.17%	Mismatches:	3
Query Match:	30.04%	Indels:	1
DB:	22	Gaps:	0

US-09-513-151-3 (1-2041) x ABB10278 (1-222)

QY	2	TGCCATAGAGGCGCTCGCTGGCGCTCACGACAGTCTCGGCGAGCGGCTCAGG	61
Db	1	CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg	20
QY	62	GGCTGCACAGGACCTACCTCTTGTAGTAGTCTCGGGCCAGCGGACCGCAAAATCC	12
Db	21	GlyLeuGlnArgThrLeuProLeuValValIleLeuGlnGlyAlaThrGlyThrGlyLysSer	40
QY	122	ACGGTGGCGTTGAGAGTAAAGCCAGCGCTCGCGCGTGAAGATCGTCAGCGCTCACTG	18
Db	41	ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValIserAlaAspSerMet	60
QY	182	CAGCTCATGAAGCGCTAGACATATCCACCAACAAGCTTCTGCCCAAGACCAAGATC	24
Db	61	GlnValIyrGlnGlyLeuAspIleIleIleThrAsnLysValIserAlaGlnGlnGlnArgIle	80
QY	242	TGCGCGACACCATGATCAGCTTTGTGGATCCTCTTGACCAATTAACAGTGGTGAC	30
Db	81	CysArgHisHisMetIleSerPheValAlaSerProLeuValThrAsnTyrThrValValAla	100
QY	302	TTGCAATAGAGCACTGCTCTGATTTGAGATATATTGGCCGACAAATTCCTATT	36
Db	101	PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle	120
QY	362	GTTGTGGAGAACCAATTTATACATTGAATCTCGCTGGAAGTCTCTGTAATACC	42
Db	121	ValValIleLys**ThrAsnTyrTyrIleGlnSerLeuLeuTrpLysValLeuValAsnThr	140
QY	422	AAGCCACGAGATGGCCACTGAGAAAGTAGTATGACCGAAAGTGCAGCTTGAAGAG	48
Db	141	LysProGlnGlu**GlyThrGlnLysValIleAspArgLysValIleGlnGlnLysGln	160
QY	482	GATGCTTTGACTTGCACAAAGCGCTAAGCAGGTGGACCCAGAAAGGCGTCCCAACTG	54
Db	161	AspGlyLeuValIleHisLysLysArgLeuSerGlnAlaIAspProGlnMetAlaAlaLysLeu	180
QY	542	CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGCTTTTGAAGAAACAGGATC	60
Db	181	HisProHisAspLysArgLysValAlaArgSerLeuGlnAlaPheGlnGlnThrGlyIle	200
QY	602	TTTCATAGTGAATTTCTCCATCGTCAACATAC-GGAAGAAAGTGGTGGTCCCT	64
Db	201	SerHisSerGlnPheLeuHisArgGlnHisThrCysLysArgTyrTrpTrpSerPro	218

RESULT 3  
ABB10476  
ID ABB10476 standard; protein; 222 AA  
XX

AC	ABBI0476:	
XX		
DT	10-JAN-2002	(first entry)
XX		
DE	Human cDNA SEQ ID NO: 784.	
XX		
KW	Human: gene therapy; neural disorder; immune system disorder;	
KW	muscular disorder; reproductive disorder; gastrointestinal disorder;	
KW	pulmonary disorder; cardiovascular disorder; renal disorder;	
XX	proliferative disorder; inflammation.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200154474-A2.	
PD	02-AUG-2001.	
XX		
PF	17-JAN-2001: 2001MO-US01349.	
XX		
PR	31-JAN-2000: 2000US-179065P.	
PR	04-FEB-2000: 2000US-180628P.	
PR	24-FEB-2000: 2000US-184664P.	
PR	02-MAR-2000: 2000US-186350P.	
PR	16-MAR-2000: 2000US-189874P.	
PR	17-MAR-2000: 2000US-190076P.	
PR	18-APR-2000: 2000US-198123P.	
PR	19-MAY-2000: 2000US-205515P.	
PR	07-JUN-2000: 2000US-209467P.	
PR	28-JUN-2000: 2000US-214866P.	
PR	30-JUN-2000: 2000US-215135P.	
PR	07-JUL-2000: 2000US-216647P.	
PR	07-JUL-2000: 2000US-216880P.	
PR	11-JUL-2000: 2000US-217487P.	
PR	11-JUL-2000: 2000US-217496P.	
PR	14-JUL-2000: 2000US-218290P.	
PR	26-JUL-2000: 2000US-220863P.	
PR	26-JUL-2000: 2000US-220964P.	
PR	14-AUG-2000: 2000US-224518P.	
PR	14-AUG-2000: 2000US-224519P.	
PR	14-AUG-2000: 2000US-225213P.	
PR	14-AUG-2000: 2000US-225214P.	
PR	14-AUG-2000: 2000US-225266P.	
PR	14-AUG-2000: 2000US-225267P.	
PR	14-AUG-2000: 2000US-225268P.	
PR	14-AUG-2000: 2000US-225270P.	
PR	14-AUG-2000: 2000US-225447P.	
PR	14-AUG-2000: 2000US-225757P.	
PR	14-AUG-2000: 2000US-225758P.	
PR	14-AUG-2000: 2000US-225759P.	
PR	18-AUG-2000: 2000US-226279P.	
PR	22-AUG-2000: 2000US-226681P.	
PR	22-AUG-2000: 2000US-226688P.	
PR	22-AUG-2000: 2000US-227182P.	
PR	23-AUG-2000: 2000US-227009P.	
PR	30-AUG-2000: 2000US-228924P.	
PR	01-SEP-2000: 2000US-229287P.	
PR	01-SEP-2000: 2000US-229343P.	
PR	01-SEP-2000: 2000US-229344P.	
PR	01-SEP-2000: 2000US-229345P.	
PR	03-SEP-2000: 2000US-229309P.	
PR	03-SEP-2000: 2000US-229513P.	
PR	06-SEP-2000: 2000US-230437P.	
PR	06-SEP-2000: 2000US-230438P.	
PR	08-SEP-2000: 2000US-231242P.	
PR	08-SEP-2000: 2000US-231243P.	
PR	08-SEP-2000: 2000US-231244P.	
PR	08-SEP-2000: 2000US-231413P.	
PR	08-SEP-2000: 2000US-231414P.	
PR	08-SEP-2000: 2000US-232080P.	
PR	12-SEP-2000: 2000US-232081P.	
PR	12-SEP-2000: 2000US-231968P.	
PR	14-SEP-2000: 2000US-232397P.	
PR	14-SEP-2000: 2000US-232398P.	

	PR	05-DEC-2000;	2000US-256719P.
	PR	06-DEC-2000;	2000US-251479P.
	PR	08-DEC-2000;	2000US-251856P.
	PR	08-DEC-2000;	2000US-251868P.
	PR	08-DEC-2000;	2000US-251869P.
	PR	08-DEC-2000;	2000US-251989P.
	PR	08-DEC-2000;	2000US-251990P.
	PR	11-DEC-2000;	2000US-254097P.
	PR	05-JAN-2001;	2001US-259678P.
XX	PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	PI	Rosen CA, Barash SC, Ruben SM;	
XX	DR	WPI: 2001-476161/51.	
XX	N-PSDB:	ABAO6698.	
PT	condition	-	
PT	Isolated nucleic acid molecule encoding an inflammation-associated polypeptide is used in preventing, treating or ameliorating a medical condition -		
PS	Claim 11; SEQ ID NO: 784; 859pp + Sequence Listing: English.		
XX	The present invention provides human cDNAs, proteins and related genomic DNAs. These can be used in the treatment of neural, immune system,		
CC	muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,		
CC	renal and proliferative disorders and inflammation. The present sequence is a protein of the invention.		
XX	Sequence 222 AA;		
SQ	Alignment Scores:		
Pred. No.:	2.68e-107	Length:	222
Score:	1074.00	Matches:	214
Percent Similarity:	98.62%	Conservative:	1
Best Local Similarity:	98.17%	Mismatches:	3
Query Match:	30.04%	Indels:	1
DB:	22	Gaps:	0
US-09-513-151-3 (1-2041) x ABB10476 (1-222)			
OY	2 TGCCTTAGATGGCGTCCGTCGGCGGTACAGAGCAGTTCCGTGCGCATGCGTGCAAG	61	
Db	1 CysHisLysMetAlaSerValAlaIalaIaaAgAlaValProValGlySerLeuArg	20	
OY	62 GGCCTGAACGAGGCCACTACCTCTTAAGTAAATTCCGGGGCACCGGACAACC	121	
Db	21 GlyleuGlInArGThrLeuProleuValIalleuGIyAlathrClYThcIylsser	40	
OY	122 AGCGTGGCGTTCACATAAGGCCAGCGGCTCGGGGTAGATCGTACAGCGTCACTCATG	181	
Db	41 ThrleuAlaleuGlInleuGIylnArGleuGIyGIurIlevaliserAlaaspserrmet	60	
OY	182 CAGGCTAATGAAGGCTGTAGACATCTGCACCAAGAAGTTTGCCCCAGACAGAGATC	241	
Db	61 GlNalTYrGlUngLYlenaspIellethrSnLySvalseAladIngIgInArGyle	80	
OY	242 TGCCGGACCAACATGATCAGCTTTGTGGATCCCTTGTGACCAATTACACAGTGTGGAC	301	
Db	81 CysArGrlSHISMelileserPhelValnsproleuValTrnAnsTyThrValValasp	100	
OY	302 TTCAGAAATAGAGCAACTGCTCTGATTTGAAGATATATTTGGCCGAGACAAAATTCCTATT	361	
Db	101 PheAdgaanArGaIatHrAlaleuIllelunspIlePheaIlaArGvasPlsyIlePrille	120	
OY	362 GTTGTGGAGAGAACCAATTATTAATCTGAACTCGCTCGGAAAGTTCCTGTCAATACC	421	
Db	121 ValValIGly***ThrAnsTyrTrllegluseerleuleutrpIysvalleuValasnThr	140	
OY	422 AAGCCCACGAGAGATGGGCACTGAGAAAGTATGACCCAAGTGGAGCTTGAAGAGAG	481	
Db	141 LysprGingluin**GlyThrGluYsvallleaAsparGlysvalGluIeuGInuYScLu	160	

Oy 482 GATGCTTTGTAAGTTCACAAAGCCCTAGACCCAGGTGACCCAGAAATGCTGCCAGCTG 541  
|||||  
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaAlaLysLeu 180  
Oy 542 CATCCACATGACAAAGCCAAAGTGCCAGACCTTGCAAGTTTGTGAAGAAACAGAAATC 601  
|||||  
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGlnThrGlyLe 200  
Oy 602 TCTCATGAGTAATTTTGTCCATGCTACACATAC-GGAGACAGCTGGTGCTCCCT 654  
|||||  
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTyrTrpSerPro 218  
RESULT 4  
AAU23398  
ID AAU23398 standard; Proteiny 222 AA.  
XX AAU23398:  
AC AAU23398:  
DT 18-DEC-2001 (first entry)  
XX  
DE Novel human enzyme polypeptide #484.  
XX  
KW Human: oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
KW autoimmune disorder; neurological disorder; metabolic disorder;  
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;  
KW nephrotoxic; anticoagulant.  
XX  
OS Homo sapiens.  
XX  
PN WO200155301-A2.  
PD 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01239.  
PF  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227100.  
PR 30-AUG-2000; 2000US-0228924.  
PR

PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR

PR 17-NOV-2000: 2000US-0249212.  
 PR 17-NOV-2000: 2000US-0249213.  
 PR 17-NOV-2000: 2000US-0249214.  
 PR 17-NOV-2000: 2000US-0249215.  
 PR 17-NOV-2000: 2000US-0249216.  
 PR 17-NOV-2000: 2000US-0249217.  
 PR 17-NOV-2000: 2000US-0249218.  
 PR 17-NOV-2000: 2000US-0249244.  
 PR 17-NOV-2000: 2000US-0249245.  
 PR 17-NOV-2000: 2000US-0249264.  
 PR 17-NOV-2000: 2000US-0249265.  
 PR 17-NOV-2000: 2000US-0249297.  
 PR 17-NOV-2000: 2000US-0249299.  
 PR 17-NOV-2000: 2000US-0249300.  
 PR 01-DEC-2000: 2000US-0250160.  
 PR 01-DEC-2000: 2000US-0250391.  
 PR 05-DEC-2000: 2000US-0251030.  
 PR 05-DEC-2000: 2000US-0251988.  
 PR 05-DEC-2000: 2000US-0256719.  
 PR 06-DEC-2000: 2000US-0251479.  
 PR 08-DEC-2000: 2000US-0251856.  
 PR 08-DEC-2000: 2000US-0251868.  
 PR 08-DEC-2000: 2000US-0251869.  
 PR 08-DEC-2000: 2000US-0251989.  
 PR 11-DEC-2000: 2000US-0254097.  
 PR 05-JAN-2001: 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM:  
 DR MPI: 2001-465566/50.  
 DR N-PSDB: AAS41268.  
 PT Novel polypeptides and polynucleotides useful for diagnosing,  
 PT preventing, treating neural, immune system, muscular, reproductive,  
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
 PT diseases  
 XX  
 PS Claim 11: SEQ ID NO 1394: 1180bp; English.  
 XX  
 CC The present invention relates to the isolation of novel human enzyme  
 CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences  
 CC encoding them. The enzyme polypeptides of the invention may comprise the  
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
 CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),  
 CC blood-related disorders (e.g. haemophilia), reproductive disorders  
 CC (e.g. infertility) and infectious disorders (e.g. influenza). The  
 CC polynucleotides of the invention can also be used in gene therapy.  
 CC AU022915-AU023814 represent the novel human enzyme polypeptides of the  
 CC invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIFO  
 CC at ftp.wifo.int/pub/published\_pcl\_sequences.  
 CC  
 XX  
 SQ Sequence 222 AA:

## Alignment Scores:

Pred. No.: 2,68e-107 Length: 222  
 Score: 1074.00 Matches: 214  
 Percent Similarity: 98.62% Conservative: 1  
 Best Local Similarity: 98.17% Mismatches: 3  
 Query Match: 30.04% Indels: 1  
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x AAU23398 (1-222)

OY 2 TGCCATAGATGGCGTCCGTCGGCCGTCGACAGACAGACTTCCTGTGGGACATGGGCTCAG 61  
 DB 1 CysHisLysMetIaSerValAlaAlaAlaArgAlaValProValIcylserGlyLeuArg 20  
 OY 62 GGCCTCAACGGACCTTACCTCTTGTAGTATTTCTCGGGCCACGGGACACGGCAATCC 121  
 DB 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40  
 OY 122 ACGCTGGCGCTTCAGCTAGGCGCCAGCGGCTCGGGGAGACATCGTCAGCCGTCCATG 181  
 DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValIserIaAspSerMet 60  
 OY 182 CAGGTCTATGAAGCGCTAGACATCATCACCACAAAGTTTCTGCCAAGACAGAGAAATC 241  
 DB 61 GlnValTyrGlnClnClnLeuAspIleIleThrAsnLysValIserIaIcnglnGlnArgIle 80  
 OY 242 TGCCGGCCACCATGATCATAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGGTGC 301  
 DB 81 CysArgHisHisMetIleSerPheValAspProLeuValIleThrAsnTyrThrValIleAsp 100  
 OY 302 TTCAGAAATAGACCAACTGCTGATTTGAAGATATATTGCCCGAGACAAATTCCTATT 361  
 DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120  
 OY 362 GTGTGGAGGAGCAATATTATTCATTGAATCTGCTGTGAAAGTTCTTGTCCATTACC 421  
 DB 121 ValValGly\*\*\*ThrAsnTyrTyrIleGluSerLeuLeuTyrLysValIleuValAsnThr 140  
 OY 422 AAGCCCCAGAGATGGGCACATGAGAAAGTATGACCGAAAGTGGACGCTTGAAGAGAG 481  
 DB 141 LysProGlnGln\*\*\*GlyThrGlnLysValIleAspArgLysValIleuGlnLysGln 160  
 OY 482 GATGCTCTGTACTTACAAACCCCTAACCGACAGCTGACCCAAATGGCTGCCAAGCTG 541  
 DB 161 AspGlyLeuValIleuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeu 180  
 OY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTCAAGTTTGTGAAGAACAGCAATC 601  
 DB 181 HisProHisAspLysArgLysValAlaIleArgSerLeuGlnValPheGlnGlnThrGlyIle 200  
 OY 602 TCTCATAGTGAATTTCTCCATCTGCAACTAC-GGAAGAAGTGGTGGTCCCT 654  
 DB 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTyrPhePro 218  
 RESULT 5  
 ID AA02526 standard; Protein: 430 AA.  
 AC AA02526;  
 XX  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE Amino acid sequence of the GRO-1 protein.  
 XX  
 KW gro-1 operon: gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;  
 KW hap-1 gene: cancer; aging; longevity; tumour formation;  
 KW physiological clock.  
 XX  
 OS Caenorhabditis elegans.  
 OS  
 PN MO9910482-A1.  
 XX  
 PD 04-MAR-1999.  
 XX  
 PF 20-AUG-1998: 98WO-CA00803.  
 XX  
 PR 25-AUG-1997: 97CA-2210251.  
 XX  
 PA (UYMC-) UNIV MCGILL.  
 XX  
 PI Barnes T, Hekimi S, Lakowski B, Lemieux J;











OY 683 TGCATCCTTGGCTTCATGCTGACACGAGCTGCTAGATGACCGCTTGATGATAGAGGGTG 742  
 Db 139 CyspHeilerpIleaspValaspIleSerValleuPhegluTyrlLeuSerLeuArgleu 158  
 OY 743 GATGACATGCTTGGCTGCTGGGCTCTTGAGGAGACTAGAGATTTTTCACAGACGCTATAT 802  
 Db 159 AspleuMetLysSerGlyMetPheclululeaglulPheHisArg----- 175  
 OY 803 CAGAGAATGTTTGGGAAATAGCCAGACATATACATGCTATCTTCCATCAATGCC 862  
 Db 176 -----SerLysLysAlaProLysgluProLeuglyIleTrpLysAlaIleclly 191  
 OY 863 TTCAAGATTTTCAGACTACTG----- 886  
 Db 192 ValGlnGluPheaspPTrLeuLysMetTyrlYstrPaspaasnaspMetasPlystrP 211  
 OY 887 -----ATCAGTGAAGGAAATGC 904  
 Db 212 AsprMetArgLysgluAlaTyrgluLysAlaValArgAlaIleLysgluAsnThrPhe 231  
 OY 905 ACAGTGGAGACTAGTACACCACTCTTAAG-----AAGGACCTGGTCCCAAT 952  
 Db 232 GlnLeu---ThrLysaspGlnIleThrLysIleasnLysLeuArgasnAlaGlyTrpasp 250  
 OY 953 GTCCCCCTGTC-----TATGGC 970  
 Db 251 IleLysLysValaspAlaThrAlaSerPheArggluAlaIleArgAlaLysgluLys 270  
 OY 971 TTAGAGTATGATGATGCTCGAG-----TGGGAGAGCTGCTTCTTGAAGCTGCTCTT 1024  
 Db 271 GluGlyValAlaIleGluMetGlnArgLysIleTrpAsnLysgluValLeuGluProCysVal 290  
 OY 1025 GAATCGTGCAGAACTTTCATC 1045  
 Db 291 LysIleValArgSerHisLeu 297  
 RESULT 10  
 AAB88507  
 ID AAB88507 standard; Protein: 311 AA.  
 AC AAB88507;  
 DT 04-JUN-2001 (first entry)  
 XX Haemophilus influenzae essential bacterial protein SED ID NO:32.  
 DE Haemophilus influenzae essential bacterial protein SED ID NO:32.  
 KW Haemophilus influenzae; essential bacterial gene; identification;  
 KW otitis media; meningitis; upper respiratory tract infection;  
 KW infection; antimicrobial.  
 XX Haemophilus influenzae.  
 OS  
 PN MO20011033-A2.  
 PD 15-FEB-2001.  
 PF 03-AUG-2000; 2000WO-US21176.  
 PR 04-AUG-1999; 99US-0368382.  
 PA (ABBO ) ABBOTT LAB.  
 P1 Chovan LE, Hessler PE, Reich KA;  
 DR WPI: 2001-147511/15.  
 DR N-PSDB: AAF94360.  
 XX  
 PT Essential bacterial genes from Haemophilus influenzae and methods for  
 PT identifying 'essential' genes that may be potential therapeutic targets  
 XX  
 PS Claim 9; Page 75-76; 185pp; English.

XX AAF94345 to AAF94409 represent essential bacterial genes from  
 CC Haemophilus influenzae, which encode the proteins given in AAB88492 to  
 CC AAB88507. The present invention also describes methods for identifying  
 CC essential bacterial genes (i.e. those essential to the survival of a  
 CC bacterium) using a transposition system. The methods are used to  
 CC identify essential genes from bacteria, especially H. influenzae (which  
 CC causes otitis media, meningitis and upper respiratory tract infections)  
 CC which may be used as targets for potential antimicrobial agents.  
 CC AAF94410 to AAF94416 represent PCR primers used in the exemplification  
 CC of the present invention.  
 XX  
 SO Sequence 311 AA;  
 Alignment Scores:  
 Pred. No: 7,96e-27 Length: 311  
 Score: 334.50 Matches: 93  
 Percent Similarity: 51.52% Conservative: 60  
 Best Local Similarity: 31.31% Mismatches: 103  
 Query Match: 9,36% Indels: 41  
 DB: Gaps: 11  
 US-09-513-151-3 (1-2041) x AAB88507 (1-311)  
 OY 86 GTAGTGAATCTCGGGCCAGCGGACCGGCAATTCAGCGCTGAGCTGACGAGCCAG 145  
 Db 6 IllePheLysMetGlyProThrAlaSerGlyLysThrAspLeuAlaIleGlnLeuArgSer 25  
 OY 146 CGGCTCGCGGCTGAGATCGTCAGCGCTGACTCCATGCGAGCTATGAGAGCCATGAGATC 205  
 Db 26 GlnLeuProValAlaGluValIleSerValAspSerAlaLeuIleTyrlLysgluMetAspIle 45  
 OY 206 ATCAACCAAGATTTTGTGCGCCAGACAGAGAATTCGCGGACCAATGATAGCTTT 265  
 Db 46 GlyThrAlaLysProSerLysgluGluLeuAlaProHisArgLeuIleAspIle 65  
 OY 266 GTGATCCTCTTGTGACCAATTAACAGATGAGTGCATTCAGAAATAGCAACTGCTCTG 325  
 Db 66 LeuAspPro---SerGluSerTyrlSerAlaMetAsnPheArgAspAlaLeuArgGlu 84  
 OY 326 ATGAGATATATTTGGCCGAGACAAATTCATGTTGTGGAGAGAACCAATATATAC 385  
 Db 85 MetAlaAspIleThrAlaGlnGlyLysIleProLeuLeuValGlyGlyThrMetLeuTyrl 104  
 OY 386 ATGAAATCTCGCTCTGGAAGTCTGTCAATCAAGCCCGAGAGATGGGACATGAG 445  
 Db 105 TyrlLysAlaLeu-----IleGlnGlyLeuSerProLeuProSerAlaAspGlu 120  
 OY 446 AAGTGAATTGACCGAAAGTGAGCTTGA-----AAGGAGATGCTCTGTA 493  
 Db 121 AsnIle-----ArgAlaGluLeuGlnGlnLysAlaAlaGlnGlnGlyTrpAlaAla 137  
 OY 494 CTTCACAAAGCCCTTACAGAGTGAGCCAGAAATGCTGCAAGCTGATCCATGAC 553  
 Db 138 LeuHisThrGluLeuAlaLysIleAspProIleSerAlaAlaArgIleAsnProSerAsp 157  
 OY 554 AAACGCAAGTGGCCAGAGCTTCAAGTTTGAAGAAACAGCAATCTCATAGTGA 613  
 Db 158 SerGlnArgIleAsnArgAlaLeuGluValPheTyrlIleThrGlyLysSerLeuThrGlu 177  
 OY 614 TTCTCCATGCTCAACATPAGGAGAGAGGTGCTCCCTGGAGAGTCTCTGAAGTTC 673  
 Db 178 Leu-----ThrGluGluLysGlyAlaLeu-----ProTyrlAspPhe 190  
 OY 674 TCTAACCTTGCATCCTTGTGCTTACATGCTGACGAGCAGCTTATGATGAGCGCTTGAT 733  
 Db 191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValLeuHisGluArgIleGlu 208  
 OY 734 AAGAGGCTGATGACATGCTTGTGCTGCGCTCTGGAGAGACTAAGAGATTTTCACAGA 793  
 Db 209 GlnArgPheHisLysMetIleGluLeuGlyPheGlnAlaGluValGluLysLeuTyrlAla 228  
 OY 794 CGCTATATATGAGAGATGTTTGGAAATAGCCAGACATATCAACATGCTATCTTCCA 853

```

Db      229 Arg-----GlyAspLeuSlnLeuProSer-----IleArg 240
      854 TCATTTGGCTTCAAGAAATTCACGAGTACTG-----
      241 CysValGlyTyrArgGlnMetTrpGluTyrLeuGlnGlyAspTyrAlaTyrGlnGluMet 260
Oy      887 ATCACTGAGGAAATGCACTGAGACTGAGTACACGCTTCTTAAGAA 937
      261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274

RESULT 11
AAU91438
ID      AAU91438 standard; Protein; 311 AA.
XX
AC      AAU91438;
XX
DT      18-JUN-2002 (first entry)
XX
DE      Haemophilus influenzae essential gene #16.
XX
KW      Essential bacterial gene; antifungal agent; antibacterial agent;
KW      antiparasitic agent; insecticidal agent; microbial infection;
KW      mucous membrane infection; otitis media; sinusitis; bronchitis;
KW      alveolitis; conjunctivitis; pneumonia; meningitis; epiglottitis;
KW      cellulitis; septic arthritis.
XX
OS      Haemophilus influenzae.
XX
PN      WO200218601-A2.
XX
PD      07-MAR-2002.
XX
PF      22-AUG-2001; 2001WO-US26245.
XX
PR      25-AUG-2000; 2000US-0649145.
XX
PA      (ABBO ) ABBOTT LAB.
XX
PI      Chovan LE, Hessler PE, Reich KA;
XX
DR      WPI: 2002-304258/34.
XX
N-PSDB: ABK64924.
XX
PT      Essential bacterial genes in Haemophilus influenzae necessary for
PT      bacterium's growth and survival, useful for screening inhibitors of
PT      polypeptides and developing therapeutic agents e.g. antimicrobial
XX
XX
PS      Claim 9: Page 75-76; 185pp; English.
XX
CC      The invention describes an essential bacterial gene (I) comprising a
CC      purified polynucleotide isolated from Haemophilus influenzae where (I)
CC      is essential to H. influenzae survival. The encoded polypeptide (II) is
CC      useful for screening substances that function to inhibit essential H.
CC      influenzae polypeptides by contacting (II) with the desired substances
CC      and measuring the response by a screen from specific, enzyme, general,
CC      affinity, phenotypic and binding screen. (I) and (II) are useful in
CC      developing therapeutic agents such as antifungal, antibacterial and
CC      antiparasitic agent, insecticidal agent, and preventive antimicrobial
CC      agents which are effective in preventing microbial infection or useful
CC      in treatment of that particular infection. (I) and (II) may also be
CC      useful in treatment of mucous membrane infections such as otitis media,
CC      sinusitis, bronchitis, alveolitis, conjunctivitis, pneumonia, meningitis,
CC      epiglottitis, cellulitis and septic arthritis. This is the amino acid
CC      sequence of an essential H. influenzae gene, described in the invention.
XX
SQ      Sequence 311 AA:

```

Alignment Scores:

Pred. No.:	7.96e-27	Length:	311
Score:	334.50	Matches:	93
Percent Similarity:	51.52%	Conservative:	60
Best Local Similarity:	31.31%	Mismatches:	103

```

Query Match: 9.36% Indels: 41
DB: 23 Gaps: 11
US-09-513-151-3 (1-2041) x AAU91438 (1-311)
Oy      86 GTAAGATATTCGGGGCCAGGGCCAGCCGCAATTCACGCTGCGCTTGCAGTACGCCAG 145
      6 IlePheLeuMetClyProThrAlaSerGlyLysThrAspLeuAlaIleGlnLeuArgSer 25
Oy      146 CGGCTGGCGCGGTAGATCGTCACGCGCTGACCTCATGCGAGGTCTATGAGCCCTAGACATC 205
      26 GlnLeuProValGlnValIleSerValIleSerValIleLeuLysGlyMetAspIle 45
Oy      206 ATGACCAACAGGTTCTGCGCCAGGACGACAGAAATGCGCGGACCAACATGACGACTT 265
      46 GlyThrAlaLysProSerLysGlnGluLeuAlaLeuAlaProHisArgLeuIleAspIle 65
Oy      266 GTGAGCTCTCTTGTGACCAATTCACAGTGTGAGTTCAGAAATGACGACTGCTG 325
      66 LeuAspPro---SerGlnSerTyrSerAlaMetAsnPheArgAspAspAlaLeuArgGlu 84
Oy      326 ATGGAATATATTTGCCCGGACAAATTCCTATGTGTGGGAGAGAACCATTTATAC 385
      85 MetAlaAspIleThrAlaGlnGlyLysIleProLeuValGlyGlyThrMetLeuTyr 104
Oy      386 ATGGAATCTGCTGTGAAAGTCTTGTCATATACCAAGCCCGAGAGATGGCAGTACG 445
      105 TyrLysAlaLeu-----IleGlnGlyLeuSerProLeuProSerAlaAspGlu 120
Oy      446 AAGATGATTGACCGAAAGTGGAGCTTGA-----AAGAGATGCTTGTGA 493
      121 AsnIle-----ArgAlaGlnLeuGlnGlnAlaIleGlnGlnGlyTrpAlaAla 137
Oy      494 CTCACAAACGCTTAAGCCGAGGAGCCGAGCCGAAATGCTGTCGCAAGCTGCATCCACATGAC 553
      138 LeuHisThrGlnLeuAlaLysIleAspProIleSerAlaIleArgIleAsnProSerAsp 157
Oy      554 AAACGCAAGTGGCCAGAGCTTGCAGAGTTTGAAGAAACAGCAATCTCTCTAGTGA 613
      158 SerGlnArgIleAsnArgAlaLeuGlnValaPheTyrIleThrGlyLysSerLeuThrGlu 177
Oy      614 TTTCCTCAGTCGACATACGAGAGAGAGTGTGTCCTTCCCTGAGGCTCTGAAATTC 673
      178 Leu-----ThrGlnGlnLysGlyAlaAlaLeu-----ProTyrAspPhe 190
Oy      674 TCATACCTTGATCTCTTGGCTTCATGCGACGACGAGCTTATAGATCGCGCTGAT 733
      191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValLeuHisGlnAlaGlnIleGlu 208
Oy      734 AAGAGGTGATGACATGCTTCTGCTGCGGCTTGTGAGAGCACTTAAGATTTTCACAGA 793
      209 GlnArgPheHisLysMetIleGlnLeuGlyPheGlnAlaGlnValaGlnLysLeuTyrAla 228
Oy      794 CGCTATTAATCAGAAATGTTTGGAAATAGCCAGGACTATACATGATGATCTTCCAA 853
      229 Arg---GlyAspLeuSlnLeuAsnLeuProSer-----IleArg 240
Oy      854 TCATTTGGCTTCAAGAAATTCACGAGTACTG-----
      241 CysValGlyTyrArgGlnMetTrpGluTyrLeuGlnGlyAspTyrAlaTyrGlnGluMet 260
Oy      887 ATCACTGAGGAAATGCACTGAGACTGAGTACACGCTTCTTAAGAA 937
      261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274

RESULT 12
AAU9765
ID      AAG49765 standard; Protein; 329 AA.
XX
AC      AAG49765;
XX
DT      18-OCT-2000 (first entry)
XX

```

DE	Arabidopsis thaliana protein fragment SEQ ID NO: 62990.	
XX	Protein identification: signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS	EP1033405-A2.	
XX	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-0301439.	
XX	25-FEB-1999; 99US-0121825.	PR 22-JUN-1999; 99US-0139899.
PR	05-MAR-1999; 99US-0123180.	PR 23-JUN-1999; 99US-0140353.
PR	09-MAR-1999; 99US-0123348.	PR 23-JUN-1999; 99US-0140354.
PR	23-MAR-1999; 99US-0125788.	PR 24-JUN-1999; 99US-0140685.
PR	25-MAR-1999; 99US-0126264.	PR 28-JUN-1999; 99US-0140823.
PR	29-MAR-1999; 99US-0126785.	PR 29-JUN-1999; 99US-0140991.
PR	01-APR-1999; 99US-0127462.	PR 30-JUN-1999; 99US-0141287.
PR	06-APR-1999; 99US-0128234.	PR 01-JUL-1999; 99US-0141842.
PR	08-APR-1999; 99US-0128714.	PR 01-JUL-1999; 99US-0142154.
PR	16-APR-1999; 99US-0129845.	PR 02-JUL-1999; 99US-0142055.
PR	19-APR-1999; 99US-0130077.	PR 06-JUL-1999; 99US-0142390.
PR	21-APR-1999; 99US-0130449.	PR 08-JUL-1999; 99US-0142803.
PR	23-APR-1999; 99US-0130510.	PR 09-JUL-1999; 99US-0142920.
PR	28-APR-1999; 99US-0130891.	PR 12-JUL-1999; 99US-0142977.
PR	30-APR-1999; 99US-0131449.	PR 13-JUL-1999; 99US-0143542.
PR	30-APR-1999; 99US-0132048.	PR 14-JUL-1999; 99US-0143624.
PR	04-MAY-1999; 99US-0132407.	PR 15-JUL-1999; 99US-0144005.
PR	05-MAY-1999; 99US-0132484.	PR 16-JUL-1999; 99US-0144085.
PR	06-MAY-1999; 99US-0132485.	PR 16-JUL-1999; 99US-0144086.
PR	06-MAY-1999; 99US-0132486.	PR 19-JUL-1999; 99US-0144325.
PR	07-MAY-1999; 99US-0132487.	PR 19-JUL-1999; 99US-0144331.
PR	11-MAY-1999; 99US-0132863.	PR 19-JUL-1999; 99US-0144332.
PR	14-MAY-1999; 99US-0134256.	PR 19-JUL-1999; 99US-0144333.
PR	14-MAY-1999; 99US-0134218.	PR 19-JUL-1999; 99US-0144334.
PR	14-MAY-1999; 99US-0134219.	PR 20-JUL-1999; 99US-0144335.
PR	14-MAY-1999; 99US-0134221.	PR 20-JUL-1999; 99US-0144352.
PR	14-MAY-1999; 99US-0134370.	PR 20-JUL-1999; 99US-0144632.
PR	18-MAY-1999; 99US-0134768.	PR 20-JUL-1999; 99US-0144884.
PR	19-MAY-1999; 99US-0134941.	PR 21-JUL-1999; 99US-0144814.
PR	20-MAY-1999; 99US-0135124.	PR 21-JUL-1999; 99US-0145086.
PR	21-MAY-1999; 99US-0135353.	PR 21-JUL-1999; 99US-0145088.
PR	24-MAY-1999; 99US-0135629.	PR 22-JUL-1999; 99US-0145085.
PR	25-MAY-1999; 99US-0136021.	PR 22-JUL-1999; 99US-0145087.
PR	27-MAY-1999; 99US-0136392.	PR 22-JUL-1999; 99US-0145089.
PR	28-MAY-1999; 99US-0136782.	PR 22-JUL-1999; 99US-0145192.
PR	01-JUN-1999; 99US-0137222.	PR 23-JUL-1999; 99US-0145145.
PR	03-JUN-1999; 99US-0137528.	PR 23-JUL-1999; 99US-0145218.
PR	04-JUN-1999; 99US-0137502.	PR 23-JUL-1999; 99US-0145224.
PR	07-JUN-1999; 99US-0137724.	PR 26-JUL-1999; 99US-0145276.
PR	08-JUN-1999; 99US-0138094.	PR 27-JUL-1999; 99US-01452913.
PR	10-JUN-1999; 99US-0138540.	PR 27-JUL-1999; 99US-0145919.
PR	14-JUN-1999; 99US-0139119.	PR 28-JUL-1999; 99US-0145951.
PR	16-JUN-1999; 99US-0139452.	PR 02-AUG-1999; 99US-0146386.
PR	16-JUN-1999; 99US-0139453.	PR 02-AUG-1999; 99US-0146388.
PR	17-JUN-1999; 99US-0139452.	PR 06-AUG-1999; 99US-0147303.
PR	18-JUN-1999; 99US-0139454.	PR 06-AUG-1999; 99US-0147316.
PR	18-JUN-1999; 99US-0139455.	PR 09-AUG-1999; 99US-0147493.
PR	18-JUN-1999; 99US-0139456.	PR 09-AUG-1999; 99US-0147935.
PR	18-JUN-1999; 99US-0139457.	PR 10-AUG-1999; 99US-0148171.
PR	18-JUN-1999; 99US-0139458.	PR 11-AUG-1999; 99US-0148319.
PR	18-JUN-1999; 99US-0139459.	PR 12-AUG-1999; 99US-0148341.
PR	18-JUN-1999; 99US-0139460.	PR 13-AUG-1999; 99US-0148565.
PR	18-JUN-1999; 99US-0139461.	PR 16-AUG-1999; 99US-0148684.
PR	18-JUN-1999; 99US-0139462.	PR 17-AUG-1999; 99US-0149175.
PR	18-JUN-1999; 99US-0139463.	PR 18-AUG-1999; 99US-0149426.
PR	18-JUN-1999; 99US-0139464.	PR 20-AUG-1999; 99US-0149722.
PR	18-JUN-1999; 99US-0139465.	PR 20-AUG-1999; 99US-0149723.
PR	18-JUN-1999; 99US-0139750.	PR 20-AUG-1999; 99US-0149929.
PR	18-JUN-1999; 99US-0139753.	PR 23-AUG-1999; 99US-0149902.
PR	21-JUN-1999; 99US-0139817.	PR 23-AUG-1999; 99US-0149930.
		PR 25-AUG-1999; 99US-0150566.
		PR 26-AUG-1999; 99US-0150884.
		PR 27-AUG-1999; 99US-0151065.
		PR 27-AUG-1999; 99US-0151066.
		PR 27-AUG-1999; 99US-0151080.

```

PR 30-AUG-1999: 99US-0151303.
PR 31-AUG-1999: 99US-0151438.
PR 01-SEP-1999: 99US-0151930.
PR 07-SEP-1999: 99US-0152363.
PR 10-SEP-1999: 99US-0153070.
PR 13-SEP-1999: 99US-0153758.
PR 15-SEP-1999: 99US-0154018.
PR 16-SEP-1999: 99US-0154039.
PR 20-SEP-1999: 99US-0154779.
PR 22-SEP-1999: 99US-0155139.
PR 23-SEP-1999: 99US-0155486.
PR 24-SEP-1999: 99US-0155659.
PR 28-SEP-1999: 99US-0156458.
PR 29-SEP-1999: 99US-0156596.
PR 04-OCT-1999: 99US-0157117.
PR 05-OCT-1999: 99US-0157753.
PR 06-OCT-1999: 99US-0157865.
PR 07-OCT-1999: 99US-0158029.
PR 08-OCT-1999: 99US-0158232.
PR 12-OCT-1999: 99US-0158369.
PR 13-OCT-1999: 99US-0159293.
PR 13-OCT-1999: 99US-0159294.
PR 13-OCT-1999: 99US-0159295.
PR 14-OCT-1999: 99US-0159329.
PR 14-OCT-1999: 99US-0159330.
PR 14-OCT-1999: 99US-0159331.
PR 14-OCT-1999: 99US-0159637.
PR 14-OCT-1999: 99US-0159638.
PR 18-OCT-1999: 99US-0159584.
PR 21-OCT-1999: 99US-0160741.
PR 21-OCT-1999: 99US-0160767.
PR 21-OCT-1999: 99US-0160768.
PR 21-OCT-1999: 99US-0160770.
PR 21-OCT-1999: 99US-0160814.
PR 21-OCT-1999: 99US-0160815.
PR 22-OCT-1999: 99US-0160980.
PR 22-OCT-1999: 99US-0160981.
PR 22-OCT-1999: 99US-0160989.
PR 25-OCT-1999: 99US-0161404.
PR 25-OCT-1999: 99US-0161405.
PR 25-OCT-1999: 99US-0161406.
PR 26-OCT-1999: 99US-0161359.
PR 26-OCT-1999: 99US-0161360.
PR 26-OCT-1999: 99US-0161361.
PR 28-OCT-1999: 99US-0161920.
PR 28-OCT-1999: 99US-0161992.
PR 28-OCT-1999: 99US-0161993.
PR 29-OCT-1999: 99US-0162142.

```

## Alignment Scores:

```

Pred. No.: 1,06e-26 Length: 329
Score: 333.50 Matches: 99
Percent Similarity: 38.948 Conservative: 56
Best Local Similarity: 24.87% Mismatches: 105
Query Match: 9.33% Indels: 138
DB: 21 Gaps: 9

```

US-09-513-151-3 (1-2041) x AAG49765 (1-329)

```

QY 83 CTGTAGTGAATTCCTCGGGCCACGGCCAGCAATTCACGCTGCGCTTGCAGCTAGCC 142
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 36 ValIlePheValMetCylValThrGlySerGlyLysSerArgLeuAlaIleAspLeuAla 55
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 143 CACGGCTCGGCGGTGATGTCACGCGCTGACTCATGCGAGCTGTATGAGCCGTAGAC 202
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 56 ThrArgPheGlnGlyIleuIleIleAsnSerAspLysIleGlnLeuTyrLysGlyLeuAsp 75
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 203 ATCATACCAACAAGGTTTCTGCCCAAGGACAGCAATTCGCGCGACCAATGATGAC 262
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 76 ValLeuThrAsnLysValThrProLysGlyCysArgGlyValProIleHisLeuLeuGly 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 263 TTGTGATGCTCTGTGACCAATTCACAGCTGTGAGCTTCGAAATAGAGCAACTGCT 322
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 96 ValIleAspSerGlyAlaGlyAsnLeuThrAlaThrGlnTyrSerArgLeuAlaSerGln 115
QY 323 CTGATTTGAAGATATATTTGCGCCGACGCAAAATTCCTATGTGTGGTGGAGCAATAT 382
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 116 AlaIleSerLysLeuSerLysAsnLysLeuProIleValAlaGlyLysSerAsnSer 135
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 383 TACATTTGAATCTCTGCTCTGGAAGTTCTTGCATACCAAGCCCGAGAGTGGCACT 442
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 136 TyrIleGluAlaLeu----- 140
QY 443 GAGAAAGTGAATTGACCGAAAGTGGAGCTTGAAAGAGGAGATGCTTGTACTTCACAAA 502
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 ----- 140
QY 503 CCGCTAAGCCAGGTGAGCCACCAAAATGGCTGCCAAGCTGCATCCACATGACAAAGCAAA 562
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 140 ----- 140
QY 563 GTGGCCAGAGCTTGCAGATTTTGAAGAAAGCAATCTCTCATAGTGAATTTCTCCAT 622
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 -----ValAsnHisSer----- 145
QY 623 CGTCAACATACGGAAGAGGTGGTGGCTTGGAGGTCCTCTGAGTTCTCTAACCT 682
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 146 -----GlyPheLeuLeuAsnTyrAspCys 154
QY 683 TGCATCTTTGGCTTCATGCTGACACGAGCTTCTGTGATGAGCGCTGATGAGAGGCT 742
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 155 CysPheIleTrpValAspValSerLeuProValLeuAsnSerPheValSerLysArgVal 174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 743 GATGACATCTGCTGCTGCTGGCTTGGAGAGCACTAAGATTTTCACAGACGCTAAT 802
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 AspArgMetMetGlnIleGlyLeuLeuGlnGluValArgGluVal-----PheAsn 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 803 CAGAGCAATGTTTGGAAATATAGCCAGACTATCAACATGATCTTCCATCAATATGCG 862
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 192 ProLys-----AlaAsnTyrSerValGlyIleArgArgAlaIleGly 205
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 863 TTCAGGAATTTTCAGAGTACCTGATCACTGAGGCAAAATGCACACTGAGACTGATAC 922
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 ValProGlnLeuHisGlnTyrLeuAsnGlnSerLeuValAspArgAlaThrLysSer 225
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 923 CAGCTTCTA-----AAGAAAGACCTGGTCCATTGTC----- 955
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 226 LysMetLeuAspValAlaValLysAsnIleLysLysAsnThrGlnIleLeuAlaCysArg 245
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 956 -----CCCCCTGTCTATGCTTAGAG 976
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 246 GlnLeuLysLysIleGlnArgLeuHisLysLysTrpLysMetSerMetHisArgValAsp 265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 977 GATCTGATGTC-----TCGAAGTGGAGAGCT 1006
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 266 AlaThrGlnValPheLeuLysArgAsnValGluGlnAlaAspGlnAlaTrpGlnAsnLeu 285
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1007 GTTCTTGAACCTGCTCTGAAATCGTCAAGTTTCATCCAGGCGCAAGCTACAGCC 1066
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 ValAlaArgProSerGlnArgIleValAspLysPheTyrAsnAsnAsn----- 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1067 ACTCAATAAAGATGCCATACAAATGAAGCTGAGACAGACAGAAATTTATACCTGTGTAC 1126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 -----AsnGlnLeuLysAsnAspAspValGlnHis-Cys----- 312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1127 CTCTGTGATGCAATCATGATGGGATGCCGAATGGCGACCCACACTAAAT 1178
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 -----LeuAlaAlaSerTyrGlyLysSerArgValAsnAsn 327
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
RESULT 13
ABB92405
ID ABB92405 standard; Protein; 329 AA.
AC ABB92405;
XX
XX
DT 31-MAY-2002 (first entry)

```

xx	Herbicideally active polypeptide SEQ ID NO 1616.
DE	Herbicide; plant; agriculture; herbicide.
xx	
KW	Arabidopsis thaliana.
xx	
OS	MO200210210-A2.
xx	
PN	07-FEB-2002.
xx	
PD	28-AUG-2001; 2001WO-EP09892.
xx	
PF	28-AUG-2001; 2001WO-EP09892.
XX	(FARB ) BAYER AG.
PA	Tietjen K, Weidler M;
xx	
PI	WP1: 2002-269010/31.
XX	
DR	
xx	Identifying plant target proteins for herbicideally active compounds,
PT	comprising aligning and comparing nucleic acid or amino acid sequences
PT	from plant with nucleic acid or amino acid sequences from non-plant
PT	organisms -
xx	
PS	Claim 5; SEQ ID NO 1616; 261pp + Sequence Listing; English.
CC	The invention relates to identifying target proteins
CC	(ABB90790-ABB94016) for herbicideally active compounds, comprising
CC	aligning and comparing nucleic acid or amino acid sequences from plant
CC	with nucleic acid or amino acid sequences from non-plant organisms using
CC	suitable search parameters, where plant sequences having an E-value
CC	greater by a factor of 3 than the E-value of most similar non-plant
CC	sequences are selected. The polypeptides or nucleic acids encoding them
CC	are useful for identifying modulators. The identified modulators are
CC	useful as herbicides.
xx	
SQ	Sequence 329 AA;
Alignment Scores:	
Pred. No.:	1,06e-26 Length: 329
Score:	333.50 Matches: 99
Percent Similarity:	38.94% Conservative: 56
Best Local Similarity:	24.87% Mismatches: 105
Query Match:	9.33% Indels: 138
DB:	23 Gaps: 9

US-09-513-151-3 (1-2041) x ABB92405 (1-329)

```

QY      83 CTTGTAGTATTCGCGGGCCACGGCAGCCGCAATCCAGCTGCCTTGACGTAGGC 142
       :::: ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     36 Valillephevalmetgylalathnclysercilysserarglenualleaspleuala 55
       :|||
QY     143 CAGCGGCTCGCGGTGATGATCGTAGCGCTGACTCCATGCAGGCTCTAGAGCCCTAGAC 202
       :||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     56 ThrtargphneglmglyguilleileansserAsphysilleglelnleuryrlsglyLeuasp 75
       :|||
QY     203 ATCATCACCAACAAGGTTTTCTGCCCAAGACAGAGAATTCGCCGCGCACCAATAGCCATGACC 262
       :::: ::::|||||:|||||:|||||:|||||:|||||:|||||:
Db     76 ValleunhrnsnlsvallthrProylselcysrfgelyalProhlshlsleulengly 95
       :|||
QY     263 TTGTGATCCTCTTGTGACCAATTACACAGTGCGTGACTTCGAATAAGCAGCACTGCT 322
       :||| ||| ||| :||| :||| :|||:
Db     96 Valpheaspsersglualaglyasnleuthralathnclotryserarlglenualasergin 115
       :|||
QY     323 CTGATTGAAGTATATTTCGCCCGACAGCAAATTCCTATTCTTGTGGAGAGAACCAATTA 382
       :||| :||| :||| :|||:|||||:|||||:|||||:
Db     116 AlaleiserlystleserAlaasnasnlyslseuProillevalalacglylserAasner 135
       :|||
QY     383 TACATTAATCTCTGCGTCTGGAAGTCTTGTCAATACCAAGCCCAGAGGATGGGCACT 442
       :|||||:|||||:|||||:|||||:|||||:|||||:
Db     136 Tyrlllegualabau----- 140
    
```

OY	443	GAGAAATGATTGACSSAAAAAGTGAGCTTTGAAAAGAGAGATGGCTTGTACTTACAA	502
Db	140	-----	140
OY	503	CGCCTAAGCCAGGTGGACCCAGAAATGGCTGCCAAGCTGCATGCATACAAAGCAAA	562
Db	140	-----	140
OY	563	GTGGCCAGAGACTTGCAGATTTTTTGAAGAACAGAAATCTCTATGATGAATTTCTGCAT	622
Db	141	-----ValAsnHisSer-----	145
OY	623	CGTCAACATACGAGAGAAGTGGTGGTCCCTTGGAGGCTCTCTGAATTTCTTAACCT	682
Db	146	-----GluPheLeuLeuAsnAsnTyrAspCys	154
OY	683	TGCATCCCTTGGCTTCATGCTGCAGACAGGAGTTCATGATGAGCGGCTTGATGAAGGGTG	742
Db	155	CysPheIleIleTyrValAspValSerLeuProValLeuAsnSerPheValSerIleCysArgVal	174
OY	743	GATGACATCTGCTGCTGCTGGCTCTTGGAGAACTAAGAGATTTTTCACAGAGCTATGAT	802
Db	175	AspArgMetMetGluAlaGlyLeuLeuGluGluValArgGluVal-----PheAsn	191
OY	803	CAGAAAGATGTTTGGGAAATTACCAGAGACTATCAACATGGTATCTTCCATCATCATGGC	862
Db	192	ProLys-----AlaAspTyrSerValGlyIleArgAlaIleGly	205
OY	863	TTCAAGGAATTTACAGAGTACCTGCATGCACAGAGGAAATGCACACTGGAGACTAGTAAC	922
Db	206	ValProGluLeuHisGluTyrLeuAspArgAsnGluSerLeuValAspArgAlaThrLysSer	225
OY	923	CAGCTTCTA-----AAGAAAGACTGGTCCATTGTC-----	955
Db	226	LysMetLeuAspValAlaValLysAsnIleLysLysAsnThrGluIleLeuValAcysArg	245
OY	956	-----CCSCCTGTATAGCTTAGAG	976
Db	246	GluLeuLysLysIleGlnArgLeuHisLysLysTrpLysMetSerMetHisArgValAsp	265
OY	977	GTAATCGATGTC-----TCAAGTGGAGAGACTCT	1006
Db	266	AlaThrGluValPheLeuLysArgAsnValGluGluAspArgValAlaIleProLysLeu	285
OY	1007	GTTCTTGAACCTGCTTGTGAATCGTGCAGAAATTTATCTACAGGCCACAAAGCTTACAGCC	1066
Db	286	ValAlaIleArgProSerGluArgIleValAspLysPheTyrAsnAsnAsn-----	301
OY	1067	ACTGCATTAAGAATGCATCATCAAGCGAGAGCGAACAAGAGAAATTAACCTGTGAC	1126
Db	302	-----AsnGlnLeuLysAsnAspAspValGluHisCys-----	312
OY	1127	CTCTGTGATCGAATCATATTGGGATCCGGAATGGGACGCCACATAAAT	1178
Db	313	-----LeuAlaIleSerTyrGluGlySerGlySerArgAlaHisAsn	327
RESULT 14			
ABB53936			
ID ABB53936 standard; Protein; 294 AA.			
XX			
AC ABB53936;			
XX			
DT 16-MAY-2002 (first entry)			
XX			
DE Lactococcus lactis protein miaa.			
XX			
KM Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.			
XX			
OS Lactococcus lactis IL1403.			
XX			
FN FR2807446-A1.			
XX			





PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139432.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140384.  
PR 24-JUN-1999; 99US-0140625.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143674.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144371.  
PR 19-JUL-1999; 99US-0144372.  
PR 19-JUL-1999; 99US-0144373.  
PR 19-JUL-1999; 99US-0144374.  
PR 19-JUL-1999; 99US-0144375.  
PR 20-JUL-1999; 99US-0144376.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151086.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159370.  
PR 14-OCT-1999; 99US-0159371.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159684.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.





Db 21 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValVal 40  
OY 239 GACCTCAGAAATAGAGCAACGCTGTGATGAGATATATATGCCCCGAGACAATTCCT 358  
Db 41 ASPPrleAAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIlePro 60  
OY 359 ATTGTGTGGAGAGCAACCAATTATACATTGATATCTGCTCTGGAAGTTCTTCTCAT 418  
Db 61 IleValValGlyGlyThrAsnTyrTyrIleGluSerLeuLeuThrLysValLeuValAsn 80  
OY 419 ACCAAGCCCCAGAGAGATGGGCACTGAGAAGTATGACCGAAAGTGGAGCTGAAAG 478  
Db 81 ThrLysProGlnGluMetGlyThrGluLysValIleAspArgLysValGluLeuGluLys 100  
OY 479 GAGATGGCTTTGTAATTCACAAAGCCTAAGCCGAGTGGACCCGGAANTGGCTCCAG 538  
Db 101 GlnAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaLys 120  
OY 539 CTGCATCCACATGACAAACGCAAGTGGCCAGAGACTTGCAAGTTTGAAGAACAAGCA 598  
Db 121 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGluThrGly 140  
OY 599 ATCTCTCATGTAATTTCTCCATCGTCACATACGGAAGAAGGTGTGCTCCCTTGA 658  
Db 141 IleSerHisSerGluPheLeuHisArgGlnHisThrGluGluGlyGlyProLeuGly 160  
OY 659 GGTCCCTCAGAGTTCTCTAACCCCTTGACATCCTTGGCTTCATCGTCAGCAGCAATCT 718  
Db 161 GlnProLeuLysPheSerAsnProCysIleLeuThrPheHisAlaAspGlnAlaValLeu 180  
OY 719 GATGAGCCCTTGATTAAGAGGTGGATGACATGCTTGGCTGGGCTCTTGAGAGACTA 778  
Db 181 AspGluArgLeuAspLysArgValAlaAspPheLeuAlaIleGlyLeuLeuGluGluLeu 200  
OY 779 AGAGATTTTCACAGACGCTATATATCAGAAGATGTTTGGAAATATAGCCAGACTATCA 838  
Db 201 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGln 220  
OY 839 CATGGTATCTCCCAATCAATGGCTTCAAGAAATTCACAGATACCTATCACTACAGGA 898  
Db 221 HisGlyIlePheGlnSerIleGlyPheLysGluPheHisGluTyrLeuIleThrGluGly 240  
OY 899 AATGACACACTGAGACAGTACAGTACAGCTTCTAAGAAGAAGA 940  
Db 241 LysCysThrLeuGluThrSerAsnGlnLeuLeuLysGly 254

RESULT 2  
US-60-453-135-7968  
: Sequence 7968, Application US/60453135  
: GENERAL INFORMATION:  
: APPLICANT: CARGILL, Michele  
: APPLICANT: IAKOUBOVA, Olga  
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
: TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
: FILE REFERENCE: C0001456  
: CURRENT APPLICATION NUMBER: US/60/453,135  
: CURRENT FILING DATE: 2003-03-10  
: NUMBER OF SEQ ID NOS: 82762  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO 7968  
: LENGTH: 475  
: TYPE: PRF  
: ORGANISM: Homo sapiens  
US-60-453-135-7968

Alignment Scores:  
Pred. No.: 3,87e-201 Length: 475  
Score: 203.00 Matches: 203  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 30.71% Indels: 0  
DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-135-7968 (1-475)  
OY 332 CATATATTTGCCCCGAGACAATAATTCCTATGCTTGGGAGAGACCAATATATACATGAA 391  
Db 116 AspIlePheAlaArgAspLysIleProIleValIleGlyGlyThrAsnTyrTyrIleGlu 135  
OY 392 TCTCGCTCTGGAAGTTCTTGTCAATACCAAGCCCGAGATGGGACATGAGAAAGT 451  
Db 136 SerLeuLeuThrLysValLeuValAlaSerHisLysProGlnGluMetGlyThrGluLysVal 155  
OY 452 ATTGACGGAAGAAGTGGAGCTTGAAGAAGAGATGGTCTTGTATCTACAAAGCCTAAGC 511  
Db 156 IleAspArgLysValGluLeuGluLysGluAspGlyLeuValLeuHisLysArgLeuSer 175  
OY 512 CAGGTGGACCCGGAATATGGCTGCCAAGTGCATCCACATGACAAACGCAAGTGGCCAG 571  
Db 176 GlnValAspProGluMetAlaIleAlaLysLeuHisProHisAspLysArgLysValAlaArg 195  
OY 572 AGCTTGCAAGTTTGAAGAACAAGATCTCTCATGTAATTTCTGCATGTCACAT 631  
Db 196 SerLeuGlnValPheGluGluThrGlyIleSerHisSerGluPheLeuHisArgGlnHis 215  
OY 632 ACGGAAGAAGTGTGTGCTCCCTTGGAGGTCTCTGAAGTTCTTAACCTTGGATCCTT 691  
Db 216 ThrGluGluGlyGlyProLeuGlyGlyProLeuLysPheSerAsnProCysIleLeu 235  
OY 692 TGGCTTCATGCTGACCCAGCAGCTTACATGAGCCCTTGATTAAGAGGTGGATACATG 751  
Db 236 ThrLeuHisAlaAspGlnAlaValLeuAspGluArgValAlaAspPhe 255  
OY 752 CTTCGTGGTGGCTTGGAGACTAAGATTTTCACAGACGCTATATATACAGAAAGAT 811  
Db 256 LeuAlaIleGlyLeuLeuGluGluLeuLeuArgAspPheHisArgArgTyrAsnGlnLysAsn 275  
OY 812 GTTTCGGAATAAGCCAGAGACTATCAACATGATCTTCCATCAATTGGCTTCAAGGAA 871  
Db 276 ValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLysGlu 295  
OY 872 TTTCCAGAGTACCTGATACAGGGAATATGACACTGAGACTAGTACCAAGCTTTGA 931  
Db 296 PheHisGluTyrLeuIleThrGluGlyLysCysThrLeuGluThrSerAsnGlnLeuLeu 315  
OY 932 AAGAAGAAGA 940  
Db 316 LysLysGly 318

RESULT 3  
US-60-453-050-7968  
: Sequence 7968, Application US/60453050  
: GENERAL INFORMATION:  
: APPLICANT: CARGILL, Michele  
: APPLICANT: LUKE, May  
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
: TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
: FILE REFERENCE: C0001457  
: CURRENT APPLICATION NUMBER: US/60/453,050  
: CURRENT FILING DATE: 2003-03-10  
: NUMBER OF SEQ ID NOS: 82762  
: SOFTWARE: FASTSEQ for Windows Version 4.0  
: SEQ ID NO 7968  
: LENGTH: 475  
: TYPE: PRF  
: ORGANISM: Homo sapiens  
US-60-453-050-7968

Alignment Scores:  
Pred. No.: 3,87e-201 Length: 475  
Score: 203.00 Matches: 203  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 30.71% Indels: 0  
DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-050-7967 (1-475)

```

OY 332 GATATATTGCCCCGAGCAAAATTCCTATTGTTGGAGAGCAACATATTACATTCGA 391
    |||||||
Db 116 AsplIeHeIaIarGAsPlYsIleProIleValIGlyIthraSTYrTylleGu 135
OY 392 TCTGTGCTGGAAAGTTTGTCAATACCAAGCCCCGAGAGGGGACACAGAAAGTG 451
    |||||||
Db 136 SerLeuEutPrLySvalLeuValaSnThrLySProGInGluMeGlyThGIduLySval 155
OY 452 ATTACCGCAAAAGTGAGACTTGAAGAGAGAGATGGTCTTGTACTTCACAAAGCCCTAGC 511
    |||||||
Db 156 IleasPrLySvalGleuGIduLySguLySguLySvalIleuHISLySargLeuSer 175
OY 512 CAGGTGAGACCCAGAAATGGCCGCAAGCTGCATCCACATGACAAACGCAAGTGGCCGAG 571
    |||||||
Db 176 GluValaSprProGluMeIaIaIaLySleuHISProHISAsPlYsargLySvalaIaIa 195
OY 572 AGCTTGCAAGTTTGAAGAAACAGAGATCTCTCATAGTGAATTTCTCCATGTCACAT 631
    |||||||
Db 196 SerLeuGInuValPheGIduThrLySleuHISSerGluPheleuHISargGInHIS 215
OY 632 AGCGAAGAGGTGGTGGTCCCTTGAGAGGCTCTGAAGTCTTAACCTTGATCTT 691
    |||||||
Db 216 ThrGluGIduLySguLySguLySguLySguLySguLySguLySguLySguLySgu 235
OY 692 TGAGCTCATGCTGACGAGCAGTTCATGATGACGCGCTTGATGATGAGGCTGATGATG 751
    |||||||
Db 236 ThrLeuHISaIaIaSprInIaIaValIleuAsPrGluArgLeuAsPrLySargLySvalaIaIa 255
OY 752 CTTCGCTGCTGGCTTGGAGGAACTTAAGAGATTTTCACAGAGGCTTAATCAGAAAT 811
    |||||||
Db 256 LeuIaIaIaGlyLeuLeuGIduLySguLySguLySguLySguLySguLySguLySgu 275
OY 812 GTTTCGGAATATGGCAGAGCTATCAACATGATCTTCCAAATCAATTTGGCTTCAGAA 871
    |||||||
Db 276 ValSerGluAsnSerGluAsPrLyGInHISGlyIlePheGInSerIleGlyPheLySgu 295
OY 872 TTTCACAGTACTGATCACTGAGGAGAAATGCACACTGGAGACTAGTAAACCACTCTTA 931
    |||||||
Db 296 PheHISgluLySguLySguLySguLySguLySguLySguLySguLySguLySgu 315
OY 932 AAGAAAGCA 940
    |||||||
Db 316 LysLySguLySgu 318

```

# RESULT 4

US-60-453-135-7967

Sequence 7967, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOUBOVA, Olga

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01456

CURRENT APPLICATION NUMBER: US/60/453.135

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7967

LENGTH: 221

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-135-7967

## Alignment Scores:

Pred. No.: 1.45e-103 Length: 221  
 Score: 109.00 Matches: 109  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 16.49% Indels: 0  
 DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-135-7967 (1-221)

```

OY 941 CCTGTCCCATTTGCCCCCTGTATAGGCTTAGAGGATCTGATGTCGAGTGGAG 1000
    |||||||
Db 91 ProGlyProIleValProProValIyrgIyLeuGIduValSerAsPrLySargTrpGlu 110
OY 1001 GAGTCTGTCTTGAACCTGCTCTTGAATCTGCAAACTTTCATCCAGGCGCACAGCCT 1060
    |||||||
Db 111 GluSerValLeuGluProIaIaIeGuIduIleValGInSerPheIleGInGlyHISLySpro 130
OY 1061 ACAGCCACTCCCAATTAAGATGCCATACATGAGCTGAGAACAGAGAGATTAACCTG 1120
    |||||||
Db 131 ThrIaIaThrProIleLyMePrGlyrAsnGluIaIaIaLySargSerTrHISLeu 150
OY 1121 TGTGACCTGTGATGCAATCTCATTTGGGAGATCGCGAATGGGACCCACATAAATCC 1180
    |||||||
Db 151 CysAsPrLeuCySAsPrArgIleIleIleGlyAsPrArgGluTrpIaIaIaHISIleLySer 170
OY 1181 AAATCCCACTTGACCAACTGAGAAAGAAAGAAAGATTTGACTCAGATGCTCAACACC 1240
    |||||||
Db 171 LysSerHISLeuAsnGInLeuLySargLySargLySargLySargLySargLySargLySarg 190
OY 1241 ATGAAAGTCAAGTGTTCGCCAGAC 1267
    |||||||
Db 191 IleGluSerGInSerValSerProAsp 199

```

## RESULT 5

US-60-453-050-7967

Sequence 7967, Application US/60453050

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: LUKE, May

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01457

CURRENT APPLICATION NUMBER: US/60/453.050

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7967

LENGTH: 221

TYPE: PRT

ORGANISM: Homo sapiens

US-60-453-050-7967

## Alignment Scores:

Pred. No.: 1.45e-103 Length: 221  
 Score: 109.00 Matches: 109  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 16.49% Indels: 0  
 DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-050-7967 (1-221)

```

OY 941 CCTGTCCCATTTGCCCCCTGTATAGGCTTAGAGGATCTGATGTCGAGTGGAG 1000
    |||||||
Db 91 ProGlyProIleValProProValIyrgIyLeuGIduValSerAsPrLySargTrpGlu 110
OY 1001 GAGTCTGTCTTGAACCTGCTCTTGAATCTGCAAACTTTCATCCAGGCGCACAGCCT 1060
    |||||||
Db 111 GluSerValLeuGluProIaIaIeGuIduIleValGInSerPheIleGInGlyHISLySpro 130
OY 1061 ACAGCCACTCCCAATTAAGATGCCATACATGAGCTGAGAACAGAGAGATTAACCTG 1120
    |||||||
Db 131 ThrIaIaThrProIleLySMePrGlyrAsnGluIaIaIaLySargSerTrHISLeu 150
OY 1121 TGTGACCTGTGATGCAATCTCATTTGGGAGATCGCGAATGGGACCCACATAAATCC 1180
    |||||||
Db 151 CysAsPrLeuCySAsPrArgIleIleIleGlyAsPrArgGluTrpIaIaIaHISIleLySer 170
OY 1181 AAATCCCACTTGACCAACTGAGAAAGAAAGAAAGATTTGACTCAGATGCTCAACACC 1240
    |||||||

```

Db 171 LysSerHisLeuAsnGlnLeuLysLysArgArgLeuAspSerAspAlaValAsnThr 190  
QY 1241 ATGAAAGTCAGACTGTTCCCGCAGC 1267  
Db 191 lIeGlSerGlnSerValSerProasp 199

## RESULT 6

US-60-453-135-7969  
Sequence 7969, Application US/60453135  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: TAKOUBOVA, Olga  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001456  
CURRENT APPLICATION NUMBER: US/60/453.135  
CURRENT FILING DATE: 2003-03-10  
NUMBER OF SEQ ID NOS: 82762  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 7969  
LENGTH: 85  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-453-135-7969

## Alignment Scores:

Pred. No.: 8,76e-56 Length: 85  
Score: 63.00 Matches: 63  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.53% Indels: 0  
DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-135-7969 (1-85)

QY 1079 ATGCCATACATGACCTGAGAACAGAGAGTATACACTGTCGTCGCTGTGATCGA 1138  
Db 1 MetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuCysAspLeuYsaAspArg 20  
QY 1139 ATCATCATTTGGGATCGCGAATGGCGCGACATTAATCCCAATCCCACTTGAACCA 1198  
Db 21 lIeIleIleGlyAspArgGlnTyrPalaAlaHisIleLysSerLysSerHisLeuAsnGln 40  
QY 1199 CTGAAGAAAGAGAGATTGAGCTCAGATGCTGCACACCATGAAGTCAGAGTGT 1258  
Db 41 LeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlSerGlnSerVal 60  
QY 1259 TCCCGCAGC 1267  
Db 61 SerProasp 63

## RESULT 7

US-60-453-050-7969  
Sequence 7969, Application US/60453050  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: LUKE, May  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001457  
CURRENT APPLICATION NUMBER: US/60/453.050  
CURRENT FILING DATE: 2003-03-10  
NUMBER OF SEQ ID NOS: 82762  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 7969  
LENGTH: 85  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-60-453-050-7969

## Alignment Scores:

Pred. No.: 8,76e-56 Length: 85  
Score: 63.00 Matches: 63

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.53% Indels: 0  
DB: 7 Gaps: 0

US-09-513-151-3 (1-2041) x US-60-453-050-7969 (1-85)

QY 1079 ATGCCATACATGACCTGAGAACAGAGAGTATACACTGTCGTCGCTGTGATCGA 1138  
Db 1 MetProTyrAsnGlnAlaGlnAsnLysArgSerTyrHisLeuCysAspLeuYsaAspArg 20  
QY 1139 ATCATCATTTGGGATCGCGAATGGCGCGACATTAATCCCAATCCCACTTGAACCA 1198  
Db 21 lIeIleIleGlyAspArgGlnTyrPalaAlaHisIleLysSerLysSerHisLeuAsnGln 40  
QY 1199 CTGAAGAAAGAGAGATTGAGCTCAGATGCTGCACACCATGAAGTCAGAGTGT 1258  
Db 41 LeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlSerGlnSerVal 60  
QY 1259 TCCCGCAGC 1267  
Db 61 SerProasp 63

## RESULT 8

US-10-203-138A-11847  
Sequence 11847, Application US/10203138A  
GENERAL INFORMATION:  
APPLICANT: Molecular Dynamics, Inc.  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL  
FILE REFERENCE: PB 0004 WO 8  
CURRENT APPLICATION NUMBER: US/10/203.138A  
PRIOR FILING DATE: 2002-08-02  
PRIOR APPLICATION NUMBER: US 60/180.312  
PRIOR FILING DATE: 04 February 2000 (04.02.00)  
PRIOR APPLICATION NUMBER: US 60/207.456  
PRIOR FILING DATE: 26 May 2000 (26.05.00)  
PRIOR APPLICATION NUMBER: US 09/632.366  
PRIOR FILING DATE: 03 August 2000 (03.08.00)  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 03 October 2000 (03.10.00)  
PRIOR APPLICATION NUMBER: US 60/236.359  
PRIOR FILING DATE: 27 September 2000 (27.09.00)  
PRIOR APPLICATION NUMBER: US 60/234.687  
PRIOR FILING DATE: 21 September 2000 (21.09.00)  
PRIOR APPLICATION NUMBER: US 09/608.408  
PRIOR FILING DATE: 30 June 2000 (30.06.00)  
NUMBER OF SEQ ID NOS: 15438  
SOFTWARE: Molecular Dynamics Sequence Listing Engine  
SEQ ID NO 11847  
LENGTH: 57  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL033527.25  
FEATURE:  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
FEATURE:  
OTHER INFORMATION: EST\_HUMAN HIT: BE242161.1, EVALUE 3.00e-21  
FEATURE:  
OTHER INFORMATION: SWISSPROT HIT: Q9KAC3, EVALUE 2.00e-04  
US-10-203-138A-11847

## Alignment Scores:

Pred. No.: 3.38e-40 Length: 57  
Score: 48.00 Matches: 48  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.26% Indels: 0

```

US-10-366-493-8141

Alignment Scores:
Pred. NO.: 7.67 Length: 365
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-366-493-8141 (1-365)

QY 14 GCGTCCGTCGGCGCTGCACGACGACTT 40
|||||
Db 85 AlaservalAlaAlaAlaArgAlaVal 93

RESULT 11
US-10-366-683-16934
; Sequence 16934, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16934
; LENGTH: 387
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-16934

Alignment Scores:
Pred. NO.: 7.61 Length: 387
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-366-683-16934 (1-387)

QY 98 GGGGCACGGGGCACCAGCAATCCACG 124
|||||
Db 135 GlyAlaThrGlyThrGlyLysSerThr 143

RESULT 12
US-10-156-761-13800
; Sequence 13800, Application US/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13800

```



```
; LENGTH: 713
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13800

Alignment Scores:
Pred. No.: 7.01 Length: 713
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-156-761-13800 (1-713)
OY 95 CTCGGGGCCAGCGGCGCAATCC 121
    |||||
Db 47 LeuGlyAlaThrGlyThrGlyLysSer 55

RESULT 13
US-10-282-122A-48736
; Sequence 48736, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48736
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48736

Alignment Scores:
Pred. No.: 96.6 Length: 124
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```

```
Query Match: 1.21% Indels: 0
DB: 6 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-282-122A-48736 (1-124)
OY 1468 TCCTTGATGCTGTTTAAGTCT 1491
    |||||
Db 57 SerPheAspValValLeuLysSer 64

RESULT 14
US-09-864-408A-124
; Sequence 124, Application US/09864408A
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: Novel Human Polynucleotides and Polypeptides Encoded There:
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-864-408A-124

Alignment Scores:
Pred. No.: 95.4 Length: 136
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 5 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-864-408A-124 (1-136)
OY 113 GGCNAATCCAGCTGGCGTTCGAC 136
    |||||
Db 14 GlyLysSerThrLeuAlaLeuGln 21

RESULT 15
US-10-366-683-18739
; Sequence 18739, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolling, Jock
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18739
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-18739

Alignment Scores:
Pred. No.: 90.6 Length: 200
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 6 Gaps: 0
```

US-09-513-151-3 (1-2041) x US-10-366-683-18739 (1-200)

Oy 14 GCGTCGTGCGGCTGCACGACA 37  
|||||  
Db 170 AlaservAlaAlaAlaAlaAlaAla 177

Search completed: April 21, 2003, 19:23:31  
Job time : 75 secs







Db 215 GlyProLeuLysPheSerAsnProCysIleLeuTrpLeuHisAlaAspGlnAlaValLeu 234  
 QY 719 GATGAGCGCTTGATAGAGGCTGATGACATGCTTGGCTGGAGGAGCTA 778  
 Db 235 AspGluArgLeuAspLysArgValAlaAspMetLeuAlaIleLeuGluGln 254  
 QY 779 AGGATTTTCACAGACGCTTAATCAAGAAGATGTTGGAAATAGCCAGACTATCA 838  
 Db 255 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGln 274  
 QY 839 CATGTAATCTCCAAATTCATTTGGCTCAAGGATTTTCAGAGTACCTGATCACTAGGGA 898  
 Db 275 HisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlnTyrLeuIleThrGlnGly 294  
 QY 899 AAATGCACACTGGAGACTAGTAACACAGCTTCTTAAGAAAGGA 940  
 Db 295 LysCysThrLeuGlnThrSerAsnGlnLeuLeuLysGly 308

## RESULT 4

US-10-097-340-151  
 ; Sequence 151, Application US/10097340  
 ; GENERAL INFORMATION:

; APPLICANT: John MONAHAN  
 ; APPLICANT: Manjula GANNAVARAPU  
 ; APPLICANT: Sebastian HOERSCH  
 ; APPLICANT: Shubhangi KAMATKAR  
 ; APPLICANT: Steve G. KOVATS  
 ; APPLICANT: Rachel E. MEYERS  
 ; APPLICANT: Michael MORRISSEY  
 ; APPLICANT: Peter OLANDT  
 ; APPLICANT: Ami SEN  
 ; APPLICANT: Peter VEIBY  
 ; APPLICANT: Gordon B. MILLS  
 ; APPLICANT: Robert C. BAST, Jr.  
 ; APPLICANT: Karen LU  
 ; APPLICANT: Rosemarie SCHMANDT  
 ; APPLICANT: Xumei ZHAO  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,  
 ; FILE REFERENCE: MRI-030  
 ; CURRENT APPLICATION NUMBER: US/10/097,340  
 ; PRIOR FILING DATE: 2002-03-14  
 ; PRIOR APPLICATION NUMBER: 60/276,025  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/325,149  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 60/276,026  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/324,967  
 ; PRIOR FILING DATE: 2001/09/26  
 ; PRIOR APPLICATION NUMBER: 60/311,732  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: 60/325,102  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 60/323,580  
 ; NUMBER OF SEQ ID NOS: 363  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 151  
 ; LENGTH: 465  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-097-340-151

## Alignment Scores:

Pred. No.: 4,1e-262 Length: 465  
 Score: 274.00 Matches: 274  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 41.45% Indels: 0  
 Db: 24 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-097-340-151 (1-465)

QY 119 TCCAGCGCTGGCTTGCAGCTAGAGCCAGCGCTGGCGGTGAGATCTGCAGCGCTACGCC 178  
 Db 35 SerThrLeuAlaLeuGlnLeuGlnLeuGlnArgLeuGlyGlyGlnIleValSerAlaAspSer 54  
 QY 179 ATCCAGAGTATGAGAGCGCTTACAGATCATCACCACAAAGCTTTGCCCAAGAGCAGAGA 238  
 Db 55 MetGlnValTyrGlnIleLeuAspIleIleThrAsnLysValSerAlaGlnGlnIleArg 74  
 QY 239 ATCTGCCGCGCACAGATGATGACCTTTGGATCCCTCTGTACCAATTAACACAGTGTG 298  
 Db 75 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValVal 94  
 QY 299 GACTTCAGAAATAGAGCAACTGCTGATTAAGATATATTGTTCCCGAGACAAATTCCT 358  
 Db 95 AspPheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIlePro 114  
 QY 359 ATTGTGTGGGAGAGAACCAATTATTAATGTAATCTGTCTGGAAAGTTCTGTCAAT 418  
 Db 115 IleValValGlyGlyThrAsnTyrTyrIleGlnSerLeuLeuTrpLysValLeuValAsn 134  
 QY 419 ACCAAGCCCGAGAGATGGCGGCTGAGAAAGATGATGACCCGAAAGTGGACTGGAAG 478  
 Db 135 ThrLysProGlnGlnMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLys 154  
 QY 479 GAGATGGCTTTGACTTACCAAAAGCGCTTAAGCCAGGAGGAGCCGGAATGGCTGCCAAG 538  
 Db 155 GluAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaValLys 174  
 QY 539 CTGCATCCATGACAAAGCGAAAGTGGCCAGAGCTTGAAGTTTGAAGAAACAGAGA 598  
 Db 175 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnThrGly 194  
 QY 599 ATCTCTAGTAATTTCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 658  
 Db 195 IleSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnGlnGlnGlnGlnGlnGln 214  
 QY 659 GGCTCTGAGAGTTCTTAACCTTGCATCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 718  
 Db 215 GlyProLeuLysPheSerAsnProCysIleLeuTrpLeuHisAlaAspGlnAlaValLeu 234  
 QY 719 GATGAGCGCTTGATAGAGGCTGATGACATGCTTGGCTGGAGGAGCTA 778  
 Db 235 AspGluArgLeuAspLysArgValAlaAspMetLeuAlaIleLeuGluGln 254  
 QY 779 AGAGATTTTCACAGACGCTTAATCAAGAAGATGTTGGAAATAGCCAGACTATCA 838  
 Db 255 ArgAspPheHisArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGln 274  
 QY 839 CATGTAATCTCCAAATTCATTTGGCTCAAGGATTTTCAGAGTACCTGATCACTAGGGA 898  
 Db 275 HisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlnTyrLeuIleThrGlnGly 294  
 QY 899 AAATGCACACTGGAGACTAGTAACACAGCTTCTTAAGAAAGGA 940  
 Db 295 LysCysThrLeuGlnThrSerAsnGlnLeuLeuLysGly 308

## RESULT 5

PCT-US01-01239-1394  
 ; Sequence 1394, Application PC/TUS0101239  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Human Genome Sciences, Inc., et al.  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 ; FILE REFERENCE: PT213PCT  
 ; CURRENT APPLICATION NUMBER: PCT/US01/01239  
 ; PRIOR FILING DATE: 2001-01-17  
 ; NUMBER OF SEQ ID NOS: 2318  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1394  
 ; LENGTH: 222  
 ; TYPE: PRT

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (124)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (145)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
PCT-US01-01239-1394

Alignment Scores:  
Pred. No.: 3,78e-112 Length: 222  
Score: 123.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 18.61% Indels: 0  
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01239-1394 (1-222)

OY 2 TGCCATAGATGGCGTCCGCGGCGGCGACGAGACAGTTCCTGTGGGCGAGTGGGCTCAGG 61  
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20  
OY 62 GGCCTGCACGACGACCTACCTCTTGTAGTATCTCCGGGCGACGGGCGACCGCAATCC 121  
Db 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40  
OY 122 ACCGTGGCGTTGACGACGACCGCGCTCGCGGTAGATCGTACGCTGACTCCATG 181  
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyIleValSerAlaAspSerMet 60  
OY 182 CAGTCTATGAGGCTGACATCATCACCACCAAGGTTTCCGCCAAGAGCAGATC 241  
Db 61 GlnValTyrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80  
OY 242 TGCCGCGACCATGATGATCCTTGTGATCCTCTTGACCAATACACAGTGGGAC 301  
Db 81 CysArgHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAlaAsp 100  
OY 302 TTCGAAATAGACGACTGCTCTGATGAGATATATTGGCCGACGACAAATTCCTATT 361  
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120  
OY 362 GTTGTGGGA 370  
Db 121 ValValGly 123

## RESULT 6

PCT-US01-01349-586  
Sequence 586, Application PC/TUS0101349

GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc., et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PJ206PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01349  
CURRENT FILING DATE: 2001-01-14  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 586  
LENGTH: 222

TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (124)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (145)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
PCT-US01-01349-586

Alignment Scores:  
Pred. No.: 3,78e-112 Length: 222  
Score: 123.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 18.61% Indels: 0  
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01349-586 (1-222)

OY 2 TGCCATAGATGGCGTCCGCGGCGGCGACGAGACAGTTCCTGTGGGCGAGTGGGCTCAGG 61  
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20  
OY 62 GGCCTGCACGACGACCTACCTCTTGTAGTATCTCCGGGCGACGGGCGACCGCAATCC 121  
Db 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40  
OY 122 ACCGTGGCGTTGACGACGACCGCGCTCGCGGTAGATCGTACGCTGACTCCATG 181  
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyIleValSerAlaAspSerMet 60  
OY 182 CAGTCTATGAGGCTGACATCATCACCACCAAGGTTTCCGCCAAGAGCAGATC 241  
Db 61 GlnValTyrGluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80  
OY 242 TGCCGCGACCATGATGATCCTTGTGATCCTCTTGACCAATACACAGTGGGAC 301  
Db 81 CysArgHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAlaAsp 100  
OY 302 TTCGAAATAGACGACTGCTCTGATGAGATATATTGGCCGACGACAAATTCCTATT 361  
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120  
OY 362 GTTGTGGGA 370  
Db 121 ValValGly 123

## RESULT 7

PCT-US01-01349-784  
Sequence 784, Application PC/TUS0101349

GENERAL INFORMATION:  
APPLICANT: Human Genome Sciences, Inc., et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PJ206PCT  
CURRENT APPLICATION NUMBER: PCT/US01/01349  
CURRENT FILING DATE: 2001-01-14  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 784  
LENGTH: 222

TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (124)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (145)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
PCT-US01-01349-784

Alignment Scores:  
Pred. No.: 3,78e-112 Length: 222  
Score: 123.00 Matches: 123  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 18.61% Indels: 0  
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01349-784 (1-222)

```

OY 2 TGCCATAGATGGCGCTCCGCTGCGCTGCAGACAGACTTCTGTGGGCGAGTGGGCTCAGG 61
    |||||||
DB 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCCTGCACGAGACCCCTACCTCTGTGATTTCTCGGGCCACGGCCAGGCAATTC 121
    |||||||
DB 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACCGTGGCGCTTGACGTAGCGCCAGCGCTGCGGTGAGATCGTCAGCGCTGCATCCATG 181
    |||||||
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAGAGCGCTTAGACATCATCACCAACAGTTTCTGCCCAAGACAGAGAAATC 241
    |||||||
DB 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerIleGlnGlnArgIle 80
OY 242 TGCCGCGACACATGATCATCTTGATCTCTGTGATCCCTGTGACCAATTAACAGTGTGAGC 301
    |||||||
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValIleThrAsnTyrThrValValAsp 100
OY 302 TTCAGAAATAGACCACTGCTGTGATTTGAAGATATATTTGCCCGAGACAAATTCCTATT 361
    |||||||
DB 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGTGGGA 370
    |||||||
DB 121 ValValGly 123

```

# RESULT 8

```

US-09-764-853-586
: Sequence 586, Application US/09764853
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PJ206
: CURRENT APPLICATION NUMBER: US/09/764,853
: PRIORITY FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 586
: LENGTH: 222
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (124)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (145)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-586

```

## Alignment Scores:

```

Pred. No.: 3 78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 21 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-764-853-586 (1-222)

```

OY 2 TGCCATAGATGGCGCTCCGCTGCGCTGCAGACAGACTTCTGTGGGCGAGTGGGCTCAGG 61
    |||||||
DB 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCCTGCACGAGACCCCTACCTCTGTGATTTCTCGGGCCACGGCCAGGCAATTC 121
    |||||||
DB 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACCGTGGCGCTTGACGTAGCGCCAGCGCTGCGGTGAGATCGTCAGCGCTGCATCCATG 181
    |||||||

```

```

DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAGAGCGCTTAGACATCATCACCAACAGTTTCTGCCCAAGACAGAGAAATC 241
    |||||||
DB 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerIleGlnGlnArgIle 80
OY 242 TGCCGCGACACATGATCATCTTGATCTCTGTGATCCCTGTGACCAATTAACAGTGTGAGC 301
    |||||||
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValIleThrAsnTyrThrValValAsp 100
OY 302 TTCAGAAATAGACCACTGCTGTGATTTGAAGATATATTTGCCCGAGACAAATTCCTATT 361
    |||||||
DB 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGTGGGA 370
    |||||||
DB 121 ValValGly 123

```

# RESULT 9

```

US-09-764-853-784
: Sequence 784, Application US/09764853
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PJ206
: CURRENT APPLICATION NUMBER: US/09/764,853
: PRIORITY FILING DATE: 2001-01-17
: Prior application data removed - consult PALM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 784
: LENGTH: 222
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (124)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (145)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-784

```

## Alignment Scores:

```

Pred. No.: 3 78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18.61% Indels: 0
DB: 21 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

```

OY 2 TGCCATAGATGGCGCTCCGCTGCGCTGCAGACAGACTTCTGTGGGCGAGTGGGCTCAGG 61
    |||||||
DB 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCCTGCACGAGACCCCTACCTCTGTGATTTCTCGGGCCACGGCCAGGCAATTC 121
    |||||||
DB 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACCGTGGCGCTTGACGTAGCGCCAGCGCTGCGGTGAGATCGTCAGCGCTGCATCCATG 181
    |||||||
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGluIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAGAGCGCTTAGACATCATCACCAACAGTTTCTGCCCAAGACAGAGAAATC 241
    |||||||
DB 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerIleGlnGlnArgIle 80
OY 242 TGCCGCGACACATGATCATCTTGATCTCTGTGATCCCTGTGACCAATTAACAGTGTGAGC 301
    |||||||
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValIleThrAsnTyrThrValValAsp 100

```



```
QY 302 TTCAGAAATAGAGCAACTGCTGATTGAGATATATTTGCCGAGCAAAATTCCTATT 361
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspIleProIle 120
QY 362 GTTGTGGGA 370
DB 121 ValValGly 123

RESULT 10
US-09-764-902-1394
; Sequence 1394, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213
; CURRENT APPLICATION NUMBER: US/09/764,902
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1394
; LENGTH: 222
; TYPE: PRP
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-902-1394

Alignment Scores:
Pred. No.: 3,78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18,61% Indels: 0
DB: 21 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-902-1394 (1-222)
QY 2 TGCCATTAAGATGCGCGTCCGCGGCGTGCACGACGACTTCTGTGGCAGCTGCGCTCAG 61
DB 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGCCTGCAAGGACCCACTCTTGTAGTATCTCGGGCCACGGGACCGGCAATCC 121
DB 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 ACCGTGCGCTTGACGTAGGCGCAGCGGCTCGCGGTGAGATCGTCAAGCGCTGACTCCATG 181
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyLulIleValSerAlaAspSerMet 60
QY 182 CAGGTCTATGAGAGCCCTAGCATCATCATCACCAAGAGTTCTGCCCAAGACGACAGATC 241
DB 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
QY 242 TGCGCGACACATGATGACGCTTGTGATCTCTTGACCAATTACAGTGTGAGAC 301
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100
QY 302 TTCAGAAATAGAGCAACTGCTGATTGAGATATATTTGCCGAGCAAAATTCCTATT 361
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspIleProIle 120
QY 362 GTTGTGGGA 370
DB 121 ValValGly 123

RESULT 11
US-10-072-326-586
```

```
; Sequence 586, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206CI
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See File wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 586
; LENGTH: 222
; TYPE: PRP
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc.feature
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-326-586

Alignment Scores:
Pred. No.: 3,78e-112 Length: 222
Score: 123.00 Matches: 123
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 18,61% Indels: 0
DB: 24 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-072-326-586 (1-222)
QY 2 TGCCATTAAGATGCGCGTCCGCGGCGTGCACGACGACTTCTGTGGCAGCTGCGCTCAG 61
DB 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGCCTGCAAGGACCCACTCTTGTAGTATCTCGGGCCACGGGACCGGCAATCC 121
DB 21 GlyLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 ACCGTGCGCTTGACGTAGGCGCAGCGGCTCGCGGTGAGATCGTCAAGCGCTGACTCCATG 181
DB 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyLulIleValSerAlaAspSerMet 60
QY 182 CAGGTCTATGAGAGCCCTAGCATCATCATCACCAAGAGTTCTGCCCAAGACGACAGATC 241
DB 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
QY 242 TGCGCGACACATGATGACGCTTGTGATCTCTTGACCAATTACAGTGTGAGAC 301
DB 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100
QY 302 TTCAGAAATAGAGCAACTGCTGATTGAGATATATTTGCCGAGCAAAATTCCTATT 361
DB 101 PheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspIleProIle 120
QY 362 GTTGTGGGA 370
DB 121 ValValGly 123

RESULT 12
US-10-072-326-784
; Sequence 784, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206CI
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See File wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1
```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 18:59:06 ; Search time 86.5 seconds

(without alignments)  
9723.518 Million cell updates/sec

Title: US-09-513-151-3

Sequence: 1 CTGCCATAGATGCGTCGCG.....TTTACAGAGAAAAA 2041

Scoring table:

OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1343044

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_n2p.model -DEV=xlh  
-Q/cg2\_1/USPTO-spool/US09513151/runat.15042003\_141246.27221/app.query.fasta.1.2183  
-DB=SPTRMBL\_21 -QFMT=fasta -SUFFIX=olindp -SPT -MIMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -STAR=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MIMLEN=0 -MAXLEN=200000000  
-USER=US09513151.ecgn.1.1.112 @runat.15042003.141246.27221 -NCPU=3  
-NO\_XLPXY -NO\_MAP -LARGEOUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7  
-Ygapop=60 -Ygapext=60 -Delop=6 -Delext=7

Database :

SPTRMBL\_21:  
1: sp\_archaea:  
2: sp\_bacteria:  
3: sp\_fungi:  
4: sp\_human:  
5: sp\_invertebrate:  
6: sp\_mammal:  
7: sp\_mhc:  
8: sp\_organelle:  
9: sp\_phage:  
10: sp\_plant:  
11: sp\_rodent:  
12: sp\_virus:  
13: sp\_vertebrate:  
14: sp\_unclassified:  
15: sp\_rvirus:  
16: sp\_bacteriap:  
17: sp\_archaeap:

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	419	63.4	435	4	Q96L45 homo sapien

2	310	46.9	467	4	Q9H3H1	Q9H3H1 homo sapien
3	169	25.6	326	4	Q9NXT7	Q9NXT7 homo sapien
4	119	18.0	324	4	Q96FJ3	Q96FJ3 homo sapien
5	46	7.0	326	11	Q9D1H5	Q9D1H5 mus musculus
6	381	1.4	382	2	Q9RLD5	Q9RLD5 pseudomonas
7	382	1.4	382	16	Q51532	Q51532 pseudomonas
8	9	1.4	434	3	Q9UT75	Q9UT75 schizosacch
9	9	1.4	434	3	Q9S9Z6	Q9S9Z6 oryza sativ
10	9	1.4	783	11	Q9QY56	Q9QY56 mus musculus
11	9	1.4	785	2	Q9F525	Q9F525 escherichia
12	9	1.2	112	16	Q9P1X9	Q9P1X9 campylobact
13	8	1.2	121	17	Q58821	Q58821 pyrococcus
14	8	1.2	131	3	Q13503	Q13503 pleurotus o
15	8	1.2	162	16	Q9RM07	Q9RM07 delinococcus
16	8	1.2	171	16	Q9KYH0	Q9KYH0 streptomyce
17	8	1.2	172	2	Q9F532	Q9F532 escherichia
18	8	1.2	173	2	Q52835	Q52835 rhizobium l
19	8	1.2	194	10	Q46323	Q46323 graciliaria
20	8	1.2	228	8	Q8WAK7	Q8WAK7 neozelebori
21	8	1.2	244	2	Q8RSL7	Q8RSL7 uncultured
22	8	1.2	246	17	Q9Y9C6	Q9Y9C6 aeropyrum p
23	8	1.2	250	11	Q924B5	Q924B5 rattus norv
24	8	1.2	263	16	Q9K798	Q9K798 bacillus ha
25	8	1.2	269	11	Q9CPQ5	Q9CPQ5 mus musculu
26	8	1.2	269	11	Q8VER8	Q8VER8 mus musculu
27	8	1.2	270	11	Q9CX70	Q9CX70 mus musculu
28	8	1.2	271	16	Q8X561	Q8X561 raltostonia s
29	8	1.2	273	17	Q8U482	Q8U482 pyrococcus
30	8	1.2	277	10	Q9SVF5	Q9SVF5 arabidopsis
31	8	1.2	289	4	Q9BX01	Q9BX01 homo sapien
32	8	1.2	301	8	Q33762	Q33762 allomyces m
33	8	1.2	312	11	Q8FV5	Q8FV5 mus musculu
34	8	1.2	320	10	Q9CA2	Q9CA2 arabidopsis
35	8	1.2	330	10	Q941D2	Q941D2 arabidopsis
36	8	1.2	330	10	Q941C9	Q941C9 arabidopsis
37	8	1.2	330	10	Q941B1	Q941B1 arabidopsis
38	8	1.2	333	10	Q8VZC9	Q8VZC9 arabidopsis
39	8	1.2	336	10	Q93WC3	Q93WC3 arabidopsis
40	8	1.2	351	5	Q9VEG7	Q9VEG7 drosophila
41	8	1.2	357	10	Q9CA35	Q9CA35 arabidopsis
42	8	1.2	357	10	Q941D3	Q941D3 arabidopsis
43	8	1.2	372	10	Q05699	Q05699 nicotiana t
44	8	1.2	374	16	Q9PB20	Q9PB20 xyella las
45	8	1.2	379	8	Q8WBMO	Q8WBMO mares zibe

## ALIGNMENTS

RESULT 1  
ID Q96L45 PRELIMINARY; PRT: 435 AA.

AC Q96L45:  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE tRNA isopentenyl transferase (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21444833; PubMed-11560893;  
RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,  
RT Barnes T., Hekimi S.,  
RT "Regulation of Physiological Rates in Caenorhabditis elegans by a  
tRNA-Modifying Enzyme in the Mitochondria.",  
RL Genetics 159:147-157(2001).  
DR EMBL: AY052768; AAL14107.1;  
DR InterPro: IPR002627; IPPT.  
DR InterPro: IPR000822; Znf\_C2H2.  
DR Pfam: PF01715; IPPT; 1.  
DR ProDom: PD004674; IPPT; 1.

DR SMART: SM00355; zNF\_C2H2; 1.  
 DR TIGR00174; miaa; 1.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 KM TRANSFERASE.  
 FT NON\_TER 1  
 SO SEQUENCE 435 AA; 48948 MW; 2279AE7C2D999FE1 CRC64;

## Alignment Scores:

Pred. No.:	0	Length:	435
Score:	419.00	Matches:	419
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	63.39%	Indels:	0
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q96L45 (1-435)

QY 29 GCACGACGATTCCTGTGGCGAGTGGGCTGACAGGCGCTGACAGGACCTGACCTTGTGA 88  
 DB 1 AlaargAlaValProValGlySerGlyLeuValrGlyLeuGlnAlaGlyThrLeuProLeuVal 20  
 QY 89 GTGATTCCTGGGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 148  
 DB 21 ValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAlaLeuGlnLeuGlyGlnArg 40  
 QY 149 CTCGGCGGTGAGATTCGTGACGCGCTGACCTGACGCTGTATGAAGCGCTGACATCATC 208  
 DB 41 LeuGlyGlyGlnIleValSerAlaAspSerMetGlnValTyrglyGlyLeuAspIleLe 60  
 QY 209 ACCAACAAGCTTTCTGCCCAAGACGACGAGCAATCTGCGGCGGCGGCGGCGGCGGCGGCTGTG 268  
 DB 61 ThrAsnLysValSerAlaGlnGlnGlnArgIleCysArgHisHisMetIleSerPheVal 80  
 QY 269 GATCCTCTGTGGACCAATTAACACAGGTGGTGAATTCAGAAATAGACAACTGCTGTGATT 328  
 DB 81 AspProLeuValThrAspTyrThrValAlaAspPheArgAsnArgAlaThrAlaLeuIle 100  
 QY 329 GAAGTATATTGGCCCGGACAAATTCCTATTGTTGGGAGGAGCAACCAATTAATTAATTT 388  
 DB 101 GlnAspIlePheAlaArgAspLysIleProIleValValGlyGlyThrAsnTyrTyrIle 120  
 QY 389 GAATCTCTGCTGTGGAAAGTCTTGTCAATACCAAGCCCGGAGATGGGCGCTGAGAA 448  
 DB 121 GluSerLeuLeuThrPlyValLeuValAsnThrLysProGlnGlnMetGlyThrGlyLys 140  
 QY 449 GTGATTCACCGAAAGTGGACCTTGAAGAAGAGATGCTGTGATTCTACAAACGCTTA 508  
 DB 141 ValIleAspArgLysValGlyLeuGlnLysGlnAspGlyLeuValIleHisLysArgLeu 160  
 QY 509 AGCCAGGCGGACCCAGAAATGGTGGCCAGCTGACATCCATGACAAACGAAAGTGGCC 568  
 DB 161 SerGlnValAspProGlnMetAlaValLysLeuHisProHisAspLysValAlaVal 180  
 QY 569 AGGAGCTTGCAGATTTTGAAGAAACAGAAATCTCTCATAGTGAATTTCTCATACGTC 628  
 DB 181 ArgSerLeuGlnValPheGlnGlnThrGlyLysSerHisSerGlnPheLeuHisArgGln 200  
 QY 629 CATACGGAAGAAGTGGTGGTCCCTTGGAGGCTCTGAAAGTTCTTAACCCCTGATC 688  
 DB 201 HisThrGlnGlnGlyGlyGlyProLeuGlyGlyProLeuLysPheSerAspProCysIle 220  
 QY 689 CTTGGGCTTCATGCTGACGACGAGTCTATGATGACGCGCTGGATGAAGAGGCTGATAC 748  
 DB 221 LeuThrPheHisAlaAspGlnAlaValLeuAspGlnAlaGlyLeuAspLysValValAsp 240  
 QY 749 ATGCTTGCCTGGGCTCTTGGAGAACTAAGATTTTCACAGACGCTAATACAGAG 808  
 DB 241 MetLeuAlaAlaGlyLeuLeuGlnGlnLysLeuValrGlyAspPheHisArgGlyTyAsnGlnLys 260  
 QY 809 AATGTTTCGGAATTAACGACGAGACTATACACATGATCTTCCATCAATCAATGGCTTCAAG 868  
 DB 261 AsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLys 280

QY 869 GAATTCACGAGTACATGATCATGAGGAAATGCACACTGGAGACTAGTAACAGCTT 928  
 DB 281 GluPheHisGlyTyrLeuIleThrGlnGlyLysCysTrpLeuGlnThrSerAsnGlnLeu 300  
 QY 929 CTAAGAAGACCTGTGTCCCATTTGTCCCGCTGTGTAGCTTATGAGAGTATGTGATGC 988  
 DB 301 LeuLysLysGlyProGlyProIleValProProValTyrGlyLeuGlnValSerAspVal 320  
 QY 969 TCGAAGTGGAGAGACTGTGTTCTTGAACCTGCTCTTGAATCGTGCAGAAATTTATCCAG 1048  
 DB 321 SerLysTrpGlnGlnSerValLeuGlnProAlaLeuGlnIleValGlnSerPheIleGln 340  
 QY 1049 GGGCAGACCGCTACAGCCGACCTCCCAATAGAATGCCATCATGGAAGCTGACAAAGAAGA 1108  
 DB 341 GlnHisLysProThrAlaThrProIleLysMetProTyrAsnGlnAlaGlnLysArg 360  
 QY 1109 AGTTATCACCTGTGTGACCTGTGTGATCGAATTCATTCATGCGGAGTCCGGAATGGGACGC 1168  
 DB 361 SerTyrHisLeuCysAspLeuCysAspArgIleIleIleGlyAspArgGlyTrpAlaAla 380  
 QY 1169 CACATAAATCCAAATCCCACTTGACCCACTGAAACCACTGAAAGAAAGAGATTGGACTCAGAT 1228  
 DB 381 HisIleLysSerLysSerHisLeuAsnGlnLeuLysLysArgArgLeuAspSerAsp 400  
 QY 1229 GCTGCAACACCATAGAAAGTCAGAGTGTTCGCCAGCTATTACAAAGACCTAAA 1285  
 DB 401 AlaValAsnThrIleGluSerGlnSerValSerProAspTyrAsnLysGlnProLys 419

## RESULT 2

Q9H3H1 PRELIMINARY; PRT; 467 AA.

AC Q9H3H1;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE tRNA isopentenylpyrophosphate transferase.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20564178; PubMed=11111046;  
 RA Golovko A., Hjalim G., Sibson F., Nicander B.;  
 RT "Cloning of a human tRNA isopentenyl transferase.";  
 RL Gene 258:85-93(2000).  
 DR EMBL: AF074918; AAC31324.1;  
 DR InterPro: IPR002627; IPT.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR Pfam: PF01715; IPT; 1.  
 DR ProDom: PD004674; IPT; 1.  
 DR TIGR00174; miaa; 1.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 KW TRANSFERASE.

US-09-513-151-3 (1-2041) x Q9H3H1 (1-467)

## Alignment Scores:

Pred. No.:	0	Length:	467
Score:	310.00	Matches:	310
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	46.90%	Indels:	0
DB:	4	Gaps:	0

QY 11 ATGCGTCGCGGCGCTGACGACGAGTTCCTGTGGCAGTGGGCTCAGGGCGCTGCAA 70  
 DB 1 MetAlaSerValAlaAlaAlaAlaArgAlaValProValGlySerGlyLeuArgGlyLeuGln 20  
 QY 71 CGGACCTTACCTCTTGTAGTGAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 130  
 DB 21 ArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40

```

OY 131 TTGCAGCTAGGCGGCGGCGGTGGATGCTGACGGCTGATGCGAGGCTAT 190
    |||||||
DB 41 LeuGlnLeuGlyGlnArgLeuGlyGlyGlnLeuValSerAlaAspSerMetGlnValTyr 60
OY 191 GAAGGCTGACATCATCCCAACAAGCTTTCTGCCCAAGACAGAGAAATCTCCCGCAC 250
    |||||||
DB 61 GluGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgLecysArgHis 80
OY 251 CACATGATCAGCTTTGTGGATCCTCTGTGACCAATTACACAGCTGGTGCATTCAGAAAT 310
    |||||||
DB 81 HisMetIleSerPheValAspProLeuValThrAsnTyrThrValAlaAspPheArgAsn 100
OY 311 AAGACAGCTGCTGATTAAGATATATTGCCCGACAAATTCCTATTGTGGGA 370
    |||||||
DB 101 ArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIleValGly 120
OY 371 GGAACCAATTATTTACATTGAATCTCTCTGGAAGTTCTTGTCAATACCAAGCCCGAC 430
    |||||||
DB 121 GlyThrAsnTyrTyrIleGlnSerLeuLeuTyrPlyValLeuValAsnThrLysProGln 140
OY 431 GAGATGGGCACTGAGAAAGTATGACCCGAAAGTGGAGCTTGAAGAGGATGGCTT 490
    |||||||
DB 141 GluMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGlnAspGlyLeu 160
OY 491 GTACTTCACAAAGCCCTAAGCCAGGTGGACCCGAAATGGCTGCAAGCTGCATCCCAT 550
    |||||||
DB 161 ValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeuHisProHis 180
OY 551 GACAAAGCGAAAGTGGCCAGAGCTTGCAGATTGTTGAAGAAACAGAGAAATCTCATAGT 610
    |||||||
DB 181 AspLysArgLysValAlaAspSerLeuGlnValPheGlnGlnThrGlyIleSerHisSer 200
OY 611 GAATTTCTCATGCTGATACATACGGAAGAAAGTGCTGCTCCCTGGAGCTCTGTAAG 670
    |||||||
DB 201 GluPheLeuHisArgGlnHisThrGlnGlnGlyGlyProLeuGlnGlyProLeuLys 220
OY 671 TTCTTAACCTTGCATCTTTGGCTTCATGCTGACAGCAGATTAGATGAGCGCTTG 730
    |||||||
DB 221 PheSerAsnProCysIleLeuTyrPheHisAlaAspIleAlaValLeuAspGlnArgLeu 240
OY 731 GATTAAGGCTGATGATGATGCTGCTGCTGGGCTCTTGAAGAACTATTTTAC 790
    |||||||
DB 241 AspLysArgValAspAspMetLeuAlaAlaGlyLeuLeuGlnGlnLeuArgAspPheHis 260
OY 791 AGACGCTATTAATCAGAAATGTTTCCGGAATAAGCCAGAGCTATCAACAGGTATCTC 850
    |||||||
DB 261 ArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePhe 280
OY 851 CAATCATTTGGCTTCAAGGAATTTACAGAGTACTGATCACTGAGGGAATAATGCACACTG 910
    |||||||
DB 281 GlnSerIleGlyPheLysGlnPheHisGlnTyrLeuIleThrGlnGlnLysCysThrLeu 300
OY 911 GAGACTAGTAACAGCTTCTTAAGAAGA 940
    |||||||
DB 301 GluThrSerAsnGlnLeuLeuLysGly 310

```

## RESULT 3

```

O9NXT7 PRELIMINARY: PRT: 326 AA.
AC O9NXT7:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CDNA FLJ20061 fls, clone COL01383.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,

```

```

RA Okitani R., Ota T., Suzuki Y., Odayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isono T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK000068; BAA90923.1;
DR InterPro: IPR002627; IPR1.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF01715; IPR1.
DR ProDom: PD004674; IPR1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2.1; UNKNOWN.1.
SQ SEQUENCE 326 AA, 37435 MW, EAA3F0F9664B7ACE CRC64:

```

## Alignment Scores:

```

Pred. No.: 3,45e-171 Length: 326
Score: 169.00 Matches: 169
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25,578 Indels: 0
DB: Gaps: 4

```

US-09-513-151-3 (1-2041) x O9NXT7 (1-326)

```

OY 434 ATGGGCACTGAGAAAGGATTTGACCCGAAAGTGGAGCTTGAAGAGAGATGCTTTGTA 493
    |||||||
DB 1 MetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGlnAspGlyLeuVal 20
OY 494 CTTCAACAAGCCCTAAGCCAGGTGGACCCGAAATGGCTGCAAGCTGCATCCCATGAC 553
    |||||||
DB 21 LeuHisLysArgLeuSerGlnValAspProGlnMetAlaAlaLysLeuHisProHisAsp 40
OY 554 AAAGCGAAAGTGGCCAGAGCTTGCAGATTGTTGAAGAAACAGAGAAATCTCATAGT 613
    |||||||
DB 41 LysArgLysValAlaAspSerLeuGlnValPheGlnGlnThrGlyIleSerHisSerGln 60
OY 614 TTCTTCATGCTGATACATACGGAAGAAAGTGCTGCTCCCTGGAGCTCTGTAAGTTC 673
    |||||||
DB 61 PheLeuHisArgGlnHisThrGlnGlnGlyGlyProLeuGlnGlyProLeuLysPhe 80
OY 674 TTCAACCTTGCATCTTTGGCTTCATGCTGACAGCAGATTAGATGAGCGCTTG 733
    |||||||
DB 81 SerAsnProCysIleLeuTyrPheHisAlaAspIleAlaValLeuAspGlnArgLeuAsp 100
OY 734 AAGAGGCTGATGATGATGCTGCTGCTGGGCTCTTGAAGAACTATTTTAC 793
    |||||||
DB 101 LysArgValAspAspMetLeuAlaAlaGlyLeuLeuGlnGlnLeuAlaGlnAspPheHisArg 120
OY 794 CGCTATTAATCAGAAATGTTTCCGGAATAAGCCAGAGCTATCAACAGGTATCTTCAA 853
    |||||||
DB 121 ArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGln 140
OY 854 TCAATTTGGCTTCAAGGAATTTACAGAGTACTGATCACTGAGGGAATAATGCACACTGAG 913
    |||||||
DB 141 SerIleGlyPheLysGlnPheHisGlnTyrLeuIleThrGlnGlnLysCysThrLeuGln 160
OY 914 ACTAGTAACAGCTTCTTAAGAAGA 940
    |||||||
DB 161 ThrSerAsnGlnLeuLeuLysGly 169

```

## RESULT 4

```

O96FJ3 PRELIMINARY: PRT: 324 AA.
AC O96FJ3:
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Similar to tRNA isopentenylpyrophosphate transferase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;

```

RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC010741; AAH10741.1;  
 DR InterPro: IPR002627; IPR1.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR Pfam: PF01715; IPR1.  
 DR ProDom: PD004674; IPR1.  
 DR SMART: SM00355; znf\_C2H2.1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 KW Transference.  
 SQ SEQUENCE 324 AA: 37223 MW: 1E6835D7C09126A9 CRC64;

## Alignment Scores:

Pred. No.:	1,13e-117	Length:	324
Score:	119.00	Matches:	167
Percent Similarity:	96.53%	Conservative:	0
Best Local Similarity:	96.53%	Mismatches:	0
Query Match:	18.00%	Indels:	6
DB:	4	Gaps:	0

US-09-513-151-3 (1-2041) x Q96FJ3 (1-324)

QY 434 ATGGCAGTGAAGTGAATGACCGAAGAGGAGCTTGAAGAGGAGGAGCTTGTGA 493  
 DB 1 MetGlyThrGluValIleAspArgLysValGluLeuGluLysGluAspGlyLeuVal 20  
 QY 494 CTTCACAAAGCCCTAAGCCAGTGGACCCAGAAATGGCTGCCAGCTGCATGCATGAC 553  
 DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeuHisProHisAsp 40  
 QY 554 AAACCCAAAGTGGCCAGAGAGCTGCAAGCTTTTGAAGAAGCAAGATCTCATAGTGA 613  
 DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyLeuSerHisSerGlu 60  
 QY 614 TTTCCTCCGTCACATACGGAAGAGCTGGTCCCTTGGAGGCTCTGCAAGTTC 673  
 DB 61 PheLeuHisArgGlnHisThrGluGluGlyGlyLeuGlyProLeuGlyGlyProLeuLysPhe 80  
 QY 674 TCTAACCTTCATCCCTTGGCTTCATGCTACACAGAGCTTCTAGATGACGGCTTGAT 733  
 DB 81 SerAsnProCysIleLeuThrPheHisAlaAspGlnAla-----AspGluArgLeuAsp 98  
 QY 734 AAGAGCGGTGATGACATGCTGCTGGCTGGCTTGGAGAACTAAGAGATTTTCACAGA 793  
 DB 99 LysArgValAspAspMetLeuAlaIleGlyLeuLeuGluGluLeuAspAspPheHisArg 118  
 QY 794 CGCTATATATCAGAGAAATGTTTGGAAATAGCAGAGACTTCAACATGATATCTTCAA 853  
 DB 119 ArgTyrAsnGlnLysAsnValSerGluAsnSerGlnAspTyrGlnHisGlyIlePheGln 138  
 QY 854 TCAATTGGCTTCAAGAAATTTACAGAGTACCTGATCAGGAGAAATGACACCTGAG 913  
 DB 139 SerIleGlyPheLysGluPheHisGlyTyrLeuIleThrGluGlyLysCysThrLeuGlu 158  
 QY 914 ACTAGTACCAAGCTTTAAAGAAAGCA 940  
 DB 159 ThrSerAsnGlnLeuLeuLysGly 167  
 RESULT 5  
 Q9DIH5 PRELIMINARY: PRT: 326 AA.  
 AC Q9DIH5;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE 2310075G14RIK protein (Riken cDNA 2310075G14 gene).  
 GN 2310075G14RIK  
 OS Mus musculus (Mouse).  
 OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:  
 OC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus:  
 OC NCBI\_TaxID=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE=EMBRYO;  
 RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavert J.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,  
 RA Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombeerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=KIDNEY;  
 RA Strausberg R.;  
 RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK003556; BAB2853.1;  
 DR EMBL: BC019812; AAH19812.1;  
 DR MGD: MGI:1914216; 2310075G14RIK.  
 DR InterPro: IPR002627; IPR1.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR Pfam: PF01715; IPR1.  
 DR ProDom: PD004674; IPR1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 SQ SEQUENCE 326 AA: 37191 MW: 454367A8B70DD1F0 CRC64;

## Alignment Scores:

Pred. No.:	1.52e-39	Length:	326
Score:	46.00	Matches:	46
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.96%	Indels:	0
DB:	11	Gaps:	0

US-09-513-151-3 (1-2041) x Q9DIH5 (1-326)

QY 494 CTTCACAAAGCCCTAAGCCAGTGGACCCAGAAATGGCTGCCAGCTGCATGCATGAC 553  
 DB 21 LeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeuHisProHisAsp 40  
 QY 554 AAACCCAAAGTGGCCAGAGAGCTGCAAGCTTTTGAAGAAGCAAGATCTCATAGTGA 613  
 DB 41 LysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyLeuSerHisSerGlu 60  
 QY 614 TTTCCTCCATGTCACAT 631  
 DB 61 PheLeuHisArgGlnHis 66  
 RESULT 6  
 Q9RLD5 PRELIMINARY: PRT: 381 AA.  
 AC Q9RLD5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Putative secretion protein.  
 GN P110.  
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).  
 OC Bacteria: Proteobacteria: gamma subdivision: Pseudomonadaceae;  
 OC Pseudomonas.  
 OC NCBI\_TaxID=316;  
 RN (1)

RP SEQUENCE FROM N.A.  
 RC STRAIN-JM300;  
 RX MEDLINE-21359310; PubMed-11466271;  
 RA Graupner S., Weger N., Schri M., Wackernagel W.;  
 RT "Requirement of bovel competence genes pilT and pilU of Pseudomonas  
 stutzeri for natural transformation and suppression of pilT deficiency  
 by a Hexahistidine tag on the type IV pilus protein PilA1.";  
 RL J. Bacteriol. 183:4694-4701(2001).  
 DR EMBL: AJ249385; CAB56296.1;  
 DR InterPro: IPR001482; GSPIL\_E.  
 DR Pfam: PF00437; GSPIL\_E.1.  
 DR ProDom: PD000739; GSPIL\_E.1.  
 SQ SEQUENCE 381 AA; 42417 MW; C69B6E6BC24EA3DB CRC64;

Alignment Scores:  
 Pred. No.: 5.94 Length: 381  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: Gaps: 0

US-09-513-151-3 (1-2041) x Q9RLD5 (1-381)  
 QY 98 GGGCCACGGCGACCGCAATCCACG 124  
 DB 130 GYAlaThrGlyThrGlyLysSerThr 138

RESULT 7  
 Q51532 PRELIMINARY; PRT; 382 AA.  
 AC 051532;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)  
 DE PilU (Twitching motility protein PilU).  
 GN PilU OR PA0396.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 15692 / PA01;  
 RX MEDLINE-91285432; PubMed-1676385;  
 RA Whittechurch C.B., Hobbs M., Livingston S.P., Krishnappillai V.,  
 RA Matlick J.S.;  
 RT "Characterisation of a Pseudomonas aeruginosa twitching motility gene  
 and evidence for a specialised protein export system widespread in  
 eubacteria.";  
 RT Gene 101:33-44(1991).  
 RL [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 15692 / PA01;  
 RA Whittechurch C.B., Matlick J.S.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 15692 / PA01;  
 RX MEDLINE-20437337; PubMed-10984043;  
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yeadu Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen.";  
 RT Nature 406:959-964(2000).  
 RL EMBL: L27667; AAA25965.1;  
 DR EMBL: AE004477; AAC03785.1;  
 DR InterPro: IPR001482; GSPIL\_E.  
 DR Pfam: PF00437; GSPIL\_E.1.

DR ProDom: PD000739; GSPIL\_E.1.  
 KW Complete Proteome.  
 SQ SEQUENCE 382 AA; 42532 MW; 02665D0FED35E17 CRC64;

Alignment Scores:  
 Pred. No.: 5.94 Length: 382  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: Gaps: 0

US-09-513-151-3 (1-2041) x Q51532 (1-381)  
 QY 98 GGGCCACGGCGACCGCAATCCACG 124  
 DB 130 GYAlaThrGlyThrGlyLysSerThr 138

RESULT 8  
 Q9UT75 PRELIMINARY; PRT; 434 AA.  
 AC 09UT75;  
 DT 01-MAY-2000 (TEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE tRNA isopentenyltransferase.  
 GN SPAC343.15.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H;  
 RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL109739; CAB52278.1;  
 DR InterPro: IPR002627; IPPT.  
 DR Pfam: PF01715; IPPT.  
 DR ProDom: PD004674; IPPT.1.  
 DR TRGFams: TIGR00174; mlaa.1.  
 KW tRNA transferase.  
 SQ SEQUENCE 434 AA; 50121 MW; FCCPD5101DF95F34D CRC64;

Alignment Scores:  
 Pred. No.: 5.84 Length: 434  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: Gaps: 0

US-09-513-151-3 (1-2041) x Q9UT75 (1-434)  
 QY 350 AAATTCCTATTCTTGTGGAGCAAC 376  
 DB 93 LysIleProIleValIleGlyGlyThr 101

RESULT 9  
 ID 08S926 PRELIMINARY; PRT; 439 AA.  
 AC 08S926;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)  
 DE Similar to tRNA isopentenyltransferase.  
 GN OJ1656.A11.14.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV: NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC
RL clone:0J1656.A11";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF003448; BAB85325.1;
KW Transferase.
SQ SEQUENCE 439 AA: 49646 MW: 74BE69EAA2EEF95 CRC64;

Alignment Scores:
Pred. No.: 5.83 Length: 439
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 10 Gaps: 0

US-09-513-151-3 (1-2041) x Q9Y56 (1-439)

QY 164 GTCAGCGCTGACTGCATGCAGCTCTAT 190
DB 64 ValserralaspsermetcInvaltyr 72

RESULT 10
Q9Y56 PRELIMINARY: PRT: 783 AA.
AC Q9Y56;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DR 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Transcription factor phi AB3.
GN E4F1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Lebedeva T.V., Singh A.K.;
RT "Repression of the murine IL-1 beta expression by the murine analog of
RL E4F transcription factor.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF126967; AAF22563.1;
DR HSSP: P15822; 32NF.
DR MGD: MGI:109530; E4F1.
DR InterPro: IPR001092; HLH_basIc.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PF00096; Zf-C2H2; 10.
DR SMART: SM00355; Znf_C2H2; 9.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 6.
DR PROSITE: PS50157; ZINC_FINGER_C2H2_2; 9.
KW DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 783 AA: 84222 MW: 41FE2E09072813PD CRC64;

Alignment Scores:
Pred. No.: 5.39 Length: 783
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 11 Gaps: 0

US-09-513-151-3 (1-2041) x Q9Y56 (1-783)

QY 998 GAGAGCTGCTGTTGTAACCTGCTCTT 1024
DB 387 GlugluserValleugluProalaleu 395

RESULT 11
Q9F525 PRELIMINARY: PRT: 785 AA.
AC Q9F525;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DR 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE TrAE protein.
GN TRAE.
OS Escherichia coli.
OC Plasmid R721.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K-12; TRANSPOSON-TN7;
RA Sampei G., Motomura K., Masuda S., Yamaguchi T., Ando K., Oishi T.,
RA Furuya N., Komano T., Mizobuchi K.;
RT "Organization and diversification of plasmid genomes: complete
RL nucleotide sequence of the R721 genome.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K-12; TRANSPOSON-TN7;
RX MEDLINE=93015772; PubMed=1400257;
RA Kim S., Komano T.;
RT "Nucleotide sequence of the R721 shuffle.";
RL J. Bacteriol. 174:7053-7058(1992).
DR EMBL: AP002527; BAB12655.1;
DR InterPro: IPR004346; Gage_TrbE_ViRB.
DR Pfam: PF03135; Gage_TrbE_ViRB; 1.
KW Plasmid.
SQ SEQUENCE 785 AA: 88299 MW: 618F5BCA7AD70D8C CRC64;

Alignment Scores:
Pred. No.: 5.39 Length: 785
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.36% Indels: 0
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x Q9F525 (1-785)

QY 92 ATTCGCGGCGCAGCGGCGCGCAA 118
DB 428 lleleuglyalarnGlyThrGlylys 436

RESULT 12
Q9PIX9 PRELIMINARY: PRT: 112 AA.
AC Q9PIX9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DR 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein Cj0148c.
GN Cj0148c.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_Taxid=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Bigham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagsels K., Kariyasev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Whitehead S., Barrett B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RL reveals hypervariable sequences.";
DR EMBL, AL139074; CAB72632.1;

```

DR InterPro: IPR003509; UPF0102.  
Pfam: PF02021; UPF0102; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 112 AA; 13179 MW; 04EEAF317625A7DE CRC64:

Alignment Scores:  
Pred. No.: 82.4 Length: 112  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.21% Indels: 0  
DB: 16 Gaps: 0

US-09-513-151-3 (1-2041) x O9P1X9 (1-112)

OY 573 GCTTGCAGTTTTCAGAGAAACAG 596  
|||||  
Db 16 AlacystylsphenulstysGln 23

RESULT 13  
O58821 PRELIMINARY: PRT: 121 AA.  
ID O58821  
AC O58821  
DT 01-AUG-1998 (TREMBlrel. 07, Created)  
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Hypothetical protein PH1094.  
GN PH1094.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_Taxid=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RX MEDLINE=98344137; Pubmed=9679194;  
RA Kawaiyayasi Y., Sawada M., Horikawa H., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Ohfuku Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudon Y., Yamazaki J., Kishida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998)  
DR EMBL: AP000004; BAA30193.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 121 AA; 12859 MW; 1BE4A036401AE523 CRC64:

Alignment Scores:  
Pred. No.: 81.5 Length: 121  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.22% Indels: 0  
DB: 17 Gaps: 0

US-09-513-151-3 (1-2041) x O58821 (1-121)

OY 1111 ACTTCTCTGTTCTCAGCTTCATT 1088  
|||||  
Db 6 ThrserrleuValleuserPheille 13

RESULT 14  
O13503 PRELIMINARY: PRT: 131 AA.  
ID O13503  
AC O13503  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE Hydrophobin precursor.  
GN POH2.  
OS Pleurotus ostreatus (oyster mushroom) (White-rot fungus).

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
OC Agaricales; Pleurotaceae; Pleurotus.  
OX NCBI\_Taxid=5322;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=FRUIT BODY;  
RA Asgelsdottilr S.A., de Vries O.M.H., Wessels J.G.H.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R7;  
RA Asgelsdottilr S.A., de Vries O.M.H., Wessels J.G.H.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Y14657; CAA74987.1;  
DR EMBL: AJ225061; CAA12392.1;  
DR InterPro: IPR001338; Hydrophobin.  
DR Pfam: PF01185; Hydrophobin; 1.  
DR SMART: SM00075; HYDRO; 1.  
DR PROSITE: PS00956; HYDROPHOBIN; 1.  
KW Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 131 POTENTIAL.  
SQ SEQUENCE 131 AA; 13114 MW; 739EA57E07FE42P1 CRC64:

Alignment Scores:  
Pred. No.: 80.6 Length: 131  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.22% Indels: 0  
DB: 3 Gaps: 0

US-09-513-151-3 (1-2041) x O13503 (1-131)

OY 774 TCCTCCAGACCCAGCAGCAACG 751  
|||||  
Db 63 SerterlySerProAlaAlaser 70

RESULT 15  
O9RM07 PRELIMINARY: PRT: 162 AA.  
ID O9RM07  
AC O9RM07  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Hypothetical protein DR0609.  
GN DR0609.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_Taxid=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R1;  
RX MEDLINE=20036896; Pubmed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
radiodurans R1.";  
RL Science 286:1571-1577(1999).  
DR EMBL: AE001919; AAF10196.1;  
DR TIGR: DR0609;  
SQ SEQUENCE 162 AA; 16942 MW; 9EE78E844F112C12 CRC64:

Alignment Scores:  
Pred. No.: 78.4 Length: 162  
Score: 8.00 Matches: 8

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	16	Gaps:	0

US-09-513-151-3 (1-2041) x Q9RM07 (1-162)

QY	107	GGCAGCGGCAATCCACGCTGGCG	130
Db	84	GlyThrGlySerThrLeuAla	91

Search completed: April 21, 2003, 19:18:35  
Job time : 94.5 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 18:55:45 ; Search time 23.5 Seconds

(without alignments)  
7204.526 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661

Sequence: 1 CTGCCTATAGATGGCGTCG.....TTTACACAGAAAAA 2041

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 112892 segs, 41476328 residues

Word size: 1

Total number of hits satisfying chosen parameters: 225614

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame\_n2p.model -DEV-xlh  
-O/cgn2.1/USPTO-spool/US09513151/runatc.15042003.141246.27211/app.query.fasta\_1.2183  
-DB=SwissProt\_40 -QPM=fastan -SUFFIX=olind2p.rsp -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=dits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=oto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09513151 -ECGN\_1\_1\_22.6runatc.15042003.141246.27211 -NCP=6 -ICU=3  
-NO\_XLPRX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-NARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	1.4	196	1	PSAD_CHLRE
2	9	1.4	428	1	PSAD_YEAST
3	9	1.4	698	1	UVRB_MYCLE
4	9	1.4	698	1	UVRB_MYCTU
5	8	1.2	92	1	RL37_EMENT
6	8	1.2	125	1	DIVC_BACSU
7	8	1.2	213	1	VARD_SULSO
8	8	1.2	216	1	KCY_CHLMU
9	8	1.2	216	1	KCY_CHLTR
10	8	1.2	216	1	KIRH_STRCO
11	8	1.2	228	1	YTOB_ERWHE
12	8	1.2	240	1	MTRB_METMA
13	8	1.2	264	1	NUDC_HAEIN
14	8	1.2	378	1	PLCD_HUMAN
15	8	1.2	417	1	YP28_CAEEL
16	8	1.2	498	1	YK04_MYCTU
17	8	1.2	580	1	NPL4_YEAST
18	8	1.2	656	1	UVRB_MYCGE

19	8	1.2	657	1	UVRB_BACCA	P56981 bacillus ca
20	8	1.2	657	1	UVRB_MYCPN	P75558 mycoplasma
21	8	1.2	658	1	UVRB_LISIN	O92844 listeria in
22	8	1.2	658	1	UVRB_LISMO	O84445 listeria mo
23	8	1.2	660	1	UVRB_BACHD	O94699 bacillus ha
24	8	1.2	661	1	UVRB_BACSU	P37954 bacillus su
25	8	1.2	663	1	UVRB_STRAA	O99417 staphylococ
26	8	1.2	663	1	UVRB_STRPY	O99425 streptococ
27	8	1.2	665	1	UVRB_ANASP	O84599 anabaena sp
28	8	1.2	665	1	UVRB_THETH	O56243 themus the
29	8	1.2	666	1	UVRB_UREPA	O9PT24 ureaplasma
30	8	1.2	669	1	UVRB_SYNY3	O55170 streptocyst
31	8	1.2	709	1	UVRB_MYCTU	P10125 micrococ
32	8	1.2	954	1	M3KA_HUMAN	O02779 homo sapien
33	8	1.2	1093	1	YK05_CAEEL	O03563 caenorhabd
34	8	1.2	1304	1	S3B1_HUMAN	O75533 homo sapien
35	8	1.2	1304	1	S3B1_MOUSE	O99409 mus musculu
36	8	1.2	1307	1	S3B1_XENLA	O57683 xenopus lae
37	8	1.2	1557	1	DVAL1_DICVI	O24702 dictyocaulu
38	7	1.1	38	1	EST5_DROMO	P10095 drosophila
39	7	1.1	62	1	SECE_SULSO	P58191 sulfolobus
40	7	1.1	98	1	ACYO_CHICK	P07032 gallus gall
41	7	1.1	98	1	YD22_MYCTU	O10635 mycobacteri
42	7	1.1	106	1	Y270_SYNY3	P73889 synechocyst
43	7	1.1	108	1	R33A_SCHPO	O94544 schizosacch
44	7	1.1	108	1	R33A_SCHPO	O94544 schizosacch
45	7	1.1	108	1	Y253_TREMA	O94595 thermotoga

#### ALIGNMENTS

RESULT 1  
PSAD\_CHLRE  
ID PSAD\_CHLRE STANDARD: PRT: 196 AA.  
AC 039615;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Photosystem I reaction center subunit II, chloroplast precursor  
DE (Photosystem I 20 kDa subunit) (PSI-D).  
GN PSAD.  
OS Chlamydomonas reinhardtii.  
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
OC Chlamydomonadaceae; Chlamydomonas.  
OX NCBI\_TaxID=3055;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95288386; PubMed=7770547;  
RA Farah J.A., Frank G., Zuber H., Rochaix J.-D.;  
RT "Cloning and sequencing of a cDNA clone encoding the photosystem I  
RT psad subunit from Chlamydomonas reinhardtii.";  
RL Plant Physiol. 107:1485-1486(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CC-406;  
RX MEDLINE=97033539; PubMed=8879236;  
RA Hahn D., Benmoun P., Kueck U.;  
RT "Altered expression of nuclear genes encoding chloroplast  
RT polypeptides in non-photosynthetic mutants of Chlamydomonas  
RT reinhardtii: evidence for post-transcriptional regulation.";  
Mol. Gen. Genet. 252:362-370(1996).  
CC - FUNCTION: PSAD CAN FORM COMPLEXES WITH FERREDOXIN AND FERREDOXIN-  
CC OXIDOREDUCTASE IN PHOTOSYSTEM I (PS I) REACTION CENTRE. PSAD MAY  
CC ENCODE THE FERREDOXIN-DOCKING PROTEIN. PSAD MAY  
CC - SUBCELLULAR LOCATION: ASSOCIATED WITH STROMAL SIDE OF THE  
CC - SIMILARITY: BELONGS TO THE PSAD FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X79674; CA56122.1; -

DR EMBL: X74419; CA52440.1; -

DR InterPro: IPR003685; Psad.

DR Pfam: PF02531; Psad; 1.

KW Photosynthesis; Photosystem I; Chloroplast; Transit peptide;

KW Thylakoid; Membrane.

FT TRANSIT 1 35 CHLOROPLAST (POTENTIAL).

FT CHAIN 36 196 PHOTOSYSTEM I REACTION CENTER SUBUNIT II.

SO SEQUENCE 196 AA; 21342 MW; 3A30F8527A896FC6 CRC64;

Alignment Scores:

Pred. No.:	1.71	Length:	196
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	1	Gaps:	0

US-09-513-151-3 (1-2041) x PSAD-CHLRE (1-196)

QY 130 GTTCAGCTAGCCAGCGGCTCGCGG 156

Db 21 VALAIAAIAArgProAlaAlaArgArg 29

RESULT 2

MOD5\_YEAST STANDARD: PRT: 428 AA.

ID MOD5\_YEAST Q12203;

AC P07884; Q12203;

DT 01-AUG-1988 (Rel. 08, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE tRNA isopentenyltransferase (EC 2.5.1.8) (isopentenyl-diphosphate: tRNA isopentenyltransferase) (IPTase) (IPTase) (IPTase).

GN MOD5 OR YOR274W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.

OX NCBI\_TaxID=4932;

RN NCBI\_TaxID=4932;

RN (1)

RP SEQUENCE FROM N.A.

RX MEDLINE=87172703; PubMed=3031457;

RA Natarian D., Dihanich M.E., Natarian N.C., Hopper A.K.;

RT "DNA sequence and transcript mapping of MOD5: features of the 5' region which suggest two translational starts.";

RL Mol. Cell. Biol. 7:185-191(1987).

RN (2)

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RX MEDLINE=97051594; PubMed=8896271;

RA Cheret G., Bernardi A., Sor F.J.;

RT "DNA sequence analysis of the VP1-SNF2 region on chromosome XV of Saccharomyces cerevisiae.";

RL Yeast 12:1059-1064(1996).

RN (3)

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE=92052176; PubMed=1946403;

RA Slusher L.B., Gillman E.C., Martin N.C., Hopper A.K.;

RT "mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene MOD5.";

RL Proc. Natl. Acad. Sci. U.S.A. 88:9789-9793(1991).

RN (4)

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE=91203856; PubMed=1850093;

RA Gillman E.C., Slusher L.B., Martin N.C., Hopper A.K.;

RT "MOD5 translation initiation sites determine N6-isopentenyladenosine modification of mitochondrial and cytoplasmic tRNA.";

RL Mol. Cell. Biol. 11:2382-2390(1991).

RN (5)

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE=94187700; PubMed=8139535;

RA Boguta M., Hunter L.A., Shen W.C., Gillman E.C., Martin N.C., Hopper A.K.;

RT "Subcellular locations of MOD5 proteins: mapping of sequences sufficient for targeting to mitochondria and demonstration that mitochondrial and nuclear isoforms commingle in the cytosol.";

RL Mol. Cell. Biol. 14:2298-2306(1994).

CC -1- FUNCTION: RESPONSIBLE FOR THE MODIFICATION OF A37 TO ISOPENTENYL A37 OF BOTH CYTOSOLIC AND MITOCHONDRIAL TRNAS.

CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.

CC -1- PATHWAY: BIOSYNTHESIS OF THE MODIFIED BASE ISOPENTENYLADENOSINE IN TRNAS.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL, CYTOPLASMIC AND NUCLEAR.

CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.

CC -1- SIMILARITY: BELONGS TO THE IPT TRANSFERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M15991; AAA34785.1; -

DR EMBL: X89633; CA61780.1; -

DR EMBL: Z75182; CA99499.1; -

DR PIR: A26717; A26717.

DR SCD: S0005800; MOD5.

DR InterPro: IPR002627; IPT.

DR Pfam: PF01715; IPTase; 1.

DR ProDom: PD004674; IPTase; 1.

DR TIGRfams: TIGR00174; miaa; 1.

DR Transferase: tRNA processing; ATP-binding; Alternative initiation; Mitochondrion; Nuclear protein.

FT FT CHAIN 1 428

FT FT CHAIN 12 428

FT FT INIT\_MET 12 12

FT FT NP\_BIND 21 28

FT FT DOMAIN 210 232

FT FT CONFLICT 313 313

FT FT CONFLICT 375 375

SO SEQUENCE 428 AA; 50236 MW; A956B17ABC05161F CRC64;

Alignment Scores:

Pred. No.:	1.54	Length:	428
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	1	Gaps:	0

US-09-513-151-3 (1-2041) x MOD5\_YEAST (1-428)

QY 350 AAAATTCCTATGTTGGAGGAACC 376

Db 104 LysleProllevalvalGlyGlyThr 112

RESULT 3

UVRB\_MYCLE STANDARD: PRT: 698 AA.

ID UVRB\_MYCLE P57991;

AC P57991;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Exonuclease ABC subunit B.

GN UVRB OR ML1387.

OS Mycobacterium leprae.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;

OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TN;  
 RX MEDLINE=21128732; PubMed=11234002;  
 RA Cole S.T., Elmgelmer K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrett B.G.;  
 RT "Massive gene decay in the leprosy bacillus."  
 RL Nature 409:1007-1011(2001).  
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT  
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS  
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES  
 CC THE APASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED  
 CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND  
 CC TO UV-IRRADIATED DUPLEX DNA (BY SIMILARITY).  
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRB, UVRB AND UVRB.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: AL583921; CAC31768.1;  
 CC DR HSSP: P56981; 1D9X.  
 CC DR Leproma: ML1387;  
 CC DR InterPro: IPR001410; DEAD.  
 CC DR InterPro: IPR001650; Helicase\_C.  
 CC DR InterPro: IPR004807; UVRB.  
 CC DR InterPro: IPR001943; UVRB/C.  
 CC DR Pfam: PF00270; DEAD. 1.  
 CC DR Pfam: PF00271; helicase\_C. 1.  
 CC DR Pfam: PF02151; UVR. 1.  
 CC DR SMART: SM00487; DEXDC. 1.  
 CC DR SMART: SM00490; HELIC\_C. 1.  
 CC DR TIGRFAMS: TIGR00631; uvrB. 1.  
 CC DR PROSITE: PS50151; UVR. 1.  
 CC KW SOS response; Excision nuclease; DNA repair; ATP-binding;  
 CC Complete proteome.  
 CC FT NP\_BIND 41 ATP (POTENTIAL).  
 CC FT DOMAIN 653 688 UVR.  
 CC SQ SEQUENCE 698 AA; 78158 MW; 7A68B02871502E6F CRC64;

Alignment Scores: 1.45 Length: 698  
 Pctd. No.: 9.00 Matches: 9  
 Score: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x UVRB\_MYCLE (1-698)

QY 95 CTCGGGCGCACGGGCGCAACATCC 121  
 ||||||||||||||||||  
 DB 40 LeuGlyAlaThrclThrglyLysSer 48

RESULT 4  
 UVRB\_MYCTU STANDARD; PRT; 698 AA.  
 ID UVRB\_MYCTU

AC 006150;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Excl nuclease ABC subunit B.  
 GN UVRB OR RV1633 OR MT1659 OR MRCY01B2.25.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Elmgelmer K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwim M.L., Haft D., Hickey E.,  
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Uitterlank T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishel W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains".  
 RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT  
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS  
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRB STIMULATES  
 CC THE APASE ACTIVITY OF UVRB IN THE PRESENCE OF UV-IRRADIATED  
 CC DOUBLE-STRANDED DNA. IT ALSO ENHANCES THE ABILITY OF UVRB TO BIND  
 CC TO UV-IRRADIATED DUPLEX DNA (BY SIMILARITY).  
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRB, UVRB AND UVRB.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE UVRB FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL: Z95554; CAB0886.1;  
 CC DR EMBL: AE007030; AAK45939.1; ALT\_INIT.  
 CC DR HSSP: P56981; 1D9X.  
 CC DR TIGR: MT1669;  
 CC DR Tuberculist: RV1633;  
 CC DR InterPro: IPR001410; DEAD.  
 CC DR InterPro: IPR001650; Helicase\_C.  
 CC DR InterPro: IPR004807; UVRB.  
 CC DR InterPro: IPR001943; UVRB/C.  
 CC DR Pfam: PF00270; DEAD. 1.  
 CC DR Pfam: PF00271; helicase\_C. 1.  
 CC DR Pfam: PF02151; UVR. 1.  
 CC DR SMART: SM00487; DEXDC. 1.  
 CC DR SMART: SM00490; HELIC\_C. 1.  
 CC DR TIGRFAMS: TIGR00631; uvrB. 1.  
 CC DR PROSITE: PS50151; UVR. 1.  
 CC KW SOS response; Excision nuclease; DNA repair; ATP-binding;  
 CC Complete proteome.

FT NP\_BIND 41 48 ATP (POTENTIAL)  
 FT DOMAIN 653 688 UVR.  
 SQ SEQUENCE 698 AA; 78070 MW; 9252A0079699CF82 CRC64;  
 Alignment Scores:  
 Score: 1.45 Length: 698  
 Percent Similarity: 9.00 Matches: 9  
 Best Local Similarity: 100.00% Conservative: 0  
 Query Match: 1.36% Mismatches: 0  
 Indels: 0  
 Gaps: 0  
 DB: 1  
 US-09-513-151-3 (1-2041) x UVRB\_MYCTU (1-698)  
 QY 95 CTGCGGGCCAGCGCCGCAATCC 121  
 DB 40 LcngAlaTnGlyThGlyLysSer 48  
 RESULT 5  
 RL37\_EMENI STANDARD: PRT: 92 AA.  
 ID RL37\_EMENI STANDARD: PRT: 92 AA.  
 AC O9COT1:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 60S ribosomal protein L37.  
 GN RPL37.  
 OS *Emicella nidulans* (*Aspergillus nidulans*).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; *Emicella*.  
 OX NCBI\_TaxID=5072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed:11179686;  
 RA Jeong H., Cho G., Han K., Kim J., Min Han D., Jahng K., Chae K.;  
 RT "Differential expression of house-keeping genes of *Aspergillus*  
 RL *nidulans* during sexual development.";  
 RL Gene 262:215-219(2001).  
 CC -1- SIMILARITY: BELONGS TO THE L37E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF277382; AK17097.1;  
 DR EMBL: AF277381; AK17096.1;  
 DR InterPro: IPR001569; Ribosomal\_L37E.  
 DR Pfam: PF01907; Ribosomal\_L37e.1.  
 DR ProDom: PD005132; Ribosomal\_L37E.1.  
 DR PROSITE: PS01077; RIBOSOMAL\_L37E; 1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 92 AA; 10502 MW; 0C924E09573DBF99 CRC64;  
 Alignment Scores:  
 Pred. No.: 21.3 Length: 92  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.22% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-513-151-3 (1-2041) x RL37\_EMENI (1-92)  
 QY 112 GTGCGCCGTGCGCCGCAATCAC 89  
 DB 85 GYAlaArgGlyProGluSnhIs 92  
 RESULT 6  
 DIVC\_BACSU

ID DIVC\_BACSU STANDARD: PRT: 125 AA.  
 AC P37471;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cell division protein divC.  
 GN DIVC OR DIVA.  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=94156852; PubMed=6113187;  
 RA Levin P.A., Losick R.;  
 RT "Characterization of a cell division gene from *Bacillus subtilis* that  
 RT is required for vegetative and sporulation septum formation.";  
 RL J. Bacteriol. 176:1451-1459(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=96051385; PubMed=7584024;  
 RA Ogasawara N., Nakai S., Yoshikawa H.;  
 RT "Systematic sequencing of the 180 kilobase region of the *Bacillus*  
 RL *subtilis* chromosome containing the replication origin.";  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borcherst S.,  
 RA Borriss R., Boutsier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Ertlen D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Filtz C., Fujita M., Fujita Y., Funa S., Gallazzi A., Galleron N.,  
 RA Gilm S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kuno M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetle D., Portolillo S., Prescott A.M.,  
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Stale Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Tempstra P., Tognoni A.,  
 RA Taseiro V., Uchiyama S., Vandenbol M., Vannier F., Vasseroiti A.,  
 RA Viari A., Wambuit R., Wedler E., Wedler H., Wiltzenegger T.,  
 RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H., Zumbstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RL *subtilis*.";  
 RL Nature 390:249-256(1997).  
 -1- FUNCTION: REQUIRED FOR VEGETATIVE AND SPORULATION SEPTUM  
 FORMATION. REQUIRED FOR THE ACTIVATION OF GENES EXPRESSED UNDER  
 THE CONTROL OF THE SPORULATION TRANSCRIPTION FACTORS SIGMA F AND  
 SIGMA E.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

CC -----
DR EMBL: L23497; AAB38379.1;
DR EMBL: D26185; BA05297.1;
DR EMBL: 299104; CAB1858.1;
DR Subtilist: BG10125; divIC.
KW Complete proteome.
SQ SEQUENCE 125 AA; 14721 MW; 2ABEB34268B5DC0 CRC64;

Alignment Scores:
Pred. No.: 20.5 Length: 125
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-513-151-3 (1-2041) x DIVC_BACSU (1-125)
OY 1193 AACCACTGAGAGAAAGAGAGA 1216
DB 24 AaagInleuLysLysArgArgArg 31

RESULT 7
VATD_SULSO STANDARD; PRT; 213 AA.
AC 09UMW9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE V-type ATP synthase subunit D (EC 3.6.3.14) (V-type ATPase subunit D).
GN APPD OR SS00566.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=2287;
RN 11
RP SEQUENCE FROM N.A.
RC SRRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=20165948; PubMed=10701121;
RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
RA Confalonieri F., Curtis B., Duguet M., Erasuo G., Faguy D.,
RA Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
RA Kishwana N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
RA Ragan M.A., Sensen C.W.;
RT "Gene content and organization of a 281-kbp contig from the genome of
the extremely thermophilic archaeon, Sulfolobus solfataricus P2."
RL Genome 43:116-136(2000).
RN 121
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=1142726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA De Moers J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2."
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
GRADIENT ACROSS THE MEMBRANE.
CC -1- CATALYTIC ACTIVITY: ATP + H(2O) + H(+)(in) -> ADP + phosphate +
H(+)(Out).
CC -1- SIMILARITY: BELONGS TO THE V-ATPASE D SUBUNIT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```

```

CC -----
CC or send an email to license@isb-sib.ch).
DR EMBL: Y18930; CAB57734.1;
DR EMBL: AE006687; AAK40881.1;
DR InterPro: IPR002699; ATPsynth_Dsub.
DR Pfam: PF01813; ATP-synth_D; 1.
DR ProDom: PD004122; ATPsynth_Dsub; 1.
DR TIGRFAMs: TIGR00309; V_ATPase_subd; 1.
KW Hydrolyase; ATP synthetase; Hydrogen ion transport; Complete proteome.
SQ SEQUENCE 213 AA; 25092 MW; 8943933520729DE8 CRC64;

Alignment Scores:
Pred. No.: 19.2 Length: 213
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: Gaps: 0

US-09-513-151-3 (1-2041) x VATD_SULSO (1-213)
OY 1858 ATTCTACAGAGAGAGAGATCA 1881
DB 203 IlleuGlnArgArgArgGluSer 210

RESULT 8
KEY_CHLUM STANDARD; PRT; 216 AA.
ID KEY_CHLUM
AC 09PJUD;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytidylate kinase (EC 2.7.4.14) (CK) (Cytidine monophosphate kinase)
DE (CMP kinase).
DE CMK OR TC0737.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN 11
RP SEQUENCE FROM N.A.
RC SRRAIN-MOPN / N199;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brumham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwin M., Nelson M., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -1- CATALYTIC ACTIVITY: ATP + (d)CMP -> ADP + (d)CDP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CYTIDYLATE KINASE FAMILY. SUBFAMILY 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC EMBL: AE002342; AAF39547.1;
CC HSSP: P23863; ICKE.
DR TIGR: TC0737;
DR InterPro: IPR003136; Cytidylylate_kin.
DR Pfam: PF02224; Cytidylylate_kin; 1.
DR TIGRFAMs: TIGR00017; cmk; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP_BIND 7
SQ SEQUENCE 216 AA; 24024 MW; 6013B4CDA7F6C59C CRC64;

Alignment Scores:

```



Pred. No.: 19.1  
 Score: 8.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 1.21%  
 DB: 1  
 Gaps: 0

US-09-513-151-3 (1-2041) x KCY\_CHLMTU (1-216)  
 QY 107 GGCACCGGCAATCCACGCGGCG 130  
 DB 10 GlyThrGlyLysSerThrLeuAla 17

RESULT 9  
 KCY\_CHLMTU  
 ID KCY\_CHLMTU STANDARD: PRT: 216 AA.  
 AC 084458:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cytidylate kinase (EC 2.7.4.14) (CK) (Cytidine monophosphate kinase)  
 DE (CMP kinase).  
 GN CMK OR CT452.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaeae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lammell C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusev R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis";  
 RL Science 282:754-759(1998).  
 CC -1- CATALYTIC ACTIVITY: ATP + (d)CMP = ADP + (d)CDP.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CYTIDYLATE KINASE FAMILY. SUBFAMILY 1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: AF001319; AAC68052.1; -  
 DR HSP, P23863; JCKE.  
 DR InterPro: IPR003136; Cytidylylate\_kin.  
 DR Pfam: PF02224; Cytidylylate\_kin.1.  
 DR TIGRFAMs: TIGR00017; cmk.1.  
 KM Transferase; Kinase; ATP-binding; Complete proteome.  
 FT NP\_BIND 7 15  
 FT ATP (By similarity).  
 SO SEQUENCE 216 AA; 24021 MW; A50CB216A9036306 CRC64;

Alignment Scores:  
 Pred. No.: 19.1  
 Score: 8.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 1.21%  
 DB: 1  
 Gaps: 0

US-09-513-151-3 (1-2041) x KCY\_CHLMTU (1-216)  
 QY 107 GGCACCGGCAATCCACGCGGCG 130  
 DB 10 GlyThrGlyLysSerThrLeuAla 17

RESULT 10  
 KITH\_STRCO

ID KITH\_STRCO STANDARD: PRT: 216 AA.  
 AC 050519:  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Thymidine kinase (EC 2.7.1.21).  
 DE TDK OR SC05845 OR SC9810.12.  
 GN Streptomyces coelicolor.  
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RC MEDLINE=21996410; PubMed=12000953;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 CC -1- CATALYTIC ACTIVITY: ATP + thymidine = ADP + thymidine 5'-  
 CC phosphate.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (potential).  
 CC -1- SIMILARITY: BELONGS TO THE THYMIDINE KINASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: AL009204; CAAL5802.1; -  
 DR InterPro: IPR001267; TK\_cell.  
 DR Pfam: PF00265; TK.1.  
 DR PROSITE: PS00603; TK\_CELLULAR\_TYPE; FALSE\_NEG.  
 KM Transferase; Kinase; DNA synthesis; ATP-binding; Complete proteome.  
 FT NP\_BIND 9 16  
 FT ATP (By similarity).  
 SO SEQUENCE 216 AA; 23368 MW; A04BA1C8C3B4ADA CRC64;

Alignment Scores:  
 Pred. No.: 19.1  
 Score: 8.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 1.21%  
 DB: 1  
 Gaps: 0

US-09-513-151-3 (1-2041) x KITH\_STRCO (1-216)  
 QY 113 GGCACCGGCAATCCACGCGGCGGCG 136  
 DB 14 GlyLysSerThrLeuAlaLeuGln 21

RESULT 11  
 YTHB\_ERWHE  
 ID YTHB\_ERWHE STANDARD: PRT: 228 AA.  
 AC 047826:  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Hypothetical protein in tute 3' region (Fragment).  
 DE Erwinia herbicola.  
 OS Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Pantoea.

OX NCBI\_TaxID=549;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 21434;  
 RA Foor F.;  
 RL Submitted (May-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE MANNITOL DEHYDROGENASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U25347; AAA6392.1;  
 DR InterPro: IPR000669; Mannitol\_dh.  
 DR Pfam: PF01232; Mannitol\_dh; 1.  
 DR PROSITE: PS00974; MANNITOL\_DEHYDROGENASE; PARTIAL.  
 KW Hypothetical protein: Oxidoreductase: NAD.  
 FT NE\_BIND 48 59 NAD (BY SIMILARITY).  
 FT NON\_TER 228 228  
 SQ SEQUENCE 228 AA; 24588 MW; 0DC5756A85FDF410 CRC64;  
 Alignment Scores:  
 Pred. No.: 19 Length: 228  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.22% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-513-151-3 (1-2041) x YTBUE\_ERWHE (1-228)  
 QY 155 GCAGGAGCCGCTGGCTAGCTGCA 132  
 ||||||||||||||||||  
 Db 193 Argarghlaalagllylealalala 200  
 RESULT 12  
 MTR\_A\_METMA STANDARD; PRT; 240 AA.  
 ID MTR\_A\_METMA  
 AC OS9640;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tetrahydromethanopterin S-methyltransferase subunit A (EC 2.1.1.86)  
 DE (N5-methyltetrahydromethanopterin--coenzyme M methyltransferase  
 DE subunit A).  
 GN MTR\_A OR M1543.  
 OS Methanosarcina mazei (Methanosarcina flisla).  
 OC Archaea: Euryarchaeota: Methanococci: Methanosarcinales;  
 OC Methanosarcinaceae: Methanosarcina.  
 OX NCBI\_TaxID=2209;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;  
 RX MEDLINE=22120827; PubMed=9559648;  
 RA Lienard T., Gottschalk G.;  
 RT Cloning, sequencing and expression of the genes encoding the sodium  
 RT translocating N5-methyltetrahydromethanopterin:coenzyme M  
 RT methyltransferase of the methylotrophic archaeon Methanosarcina mazei  
 RT GoI.  
 RL FEBS Lett. 425:204-208(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;  
 RX MEDLINE=22120827; PubMed=12125824;  
 RA Deppeleier U., Johann A., Hartsch T., Merl R., Schmitz R.A.,  
 RA Matlener A., Lienard T., Heme A., Wier A., Baumer S., Jacobi C.,  
 RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,  
 RA Blatshcharyya A., Lykdis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
 RA Fritz H.-J., Gottschalk G.;

RT "The genome of Methanosarcina mazei: evidence for lateral gene  
 RT transfer between Bacteria and Archaea."  
 RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
 CC -1- FUNCTION: THIS ENZYME COMPLEX CATALYZES AN INTERMEDIATE STEP IN  
 CC METHANOGENESIS, THE FORMATION OF METHYL-COENZYME M AND  
 CC TETRAHYDROMETHANOPTERIN FROM COENZYME M AND N5-METHYL-  
 CC TETRAHYDROMETHANOPTERIN. THIS IS A ENERGY-CONSERVING, SODIUM-ION  
 CC TRANSLOCATING STEP.  
 CC -1- CATALYTIC ACTIVITY: 5-methyl-5,6,7,8-tetrahydromethanopterin + 2-  
 CC mercaptoethanesulfonate = 5,6,7,8-tetrahydromethanopterin + 2-  
 CC (methylthio)ethanesulfonate.  
 CC -1- COFACTOR: BINDS 5-HYDROXYBENZIMIDAZOL COBAMIDE AS A PROSTHETIC  
 CC GROUP.  
 CC -1- PATHWAY: Methanogenesis.  
 CC -1- SUBUNIT: COMPOSED OF 8 DIFFERENT SUBUNITS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND VESICULAR MEMBRANE-  
 CC ASSOCIATED (BY SIMILARITY).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF042381; AAC38334.1;  
 DR EMBL: AE013388; AAM31239.1;  
 DR TIGRfams: TIGR01111; mtrA; 1.  
 KW Transferase: Methyltransferase; Methanogenesis; Cobalt.  
 FT BINDING 85 85 5-HYDROXYBENZIMIDAZOL COBAMIDE COFACTOR  
 (BY SIMILARITY).  
 FT DOMAIN 169 176 POLY-GLU.  
 FT DOMAIN 232 238 POLY-LEU.  
 FT CONFLICT 16 16 G -> R (IN REF. 1).  
 FT CONFLICT 71 71 N -> D (IN REF. 1).  
 FT CONFLICT 127 127 E -> G (IN REF. 1).  
 FT CONFLICT 161 161 D -> E (IN REF. 1).  
 SQ SEQUENCE 240 AA; 25366 MW; 44C086DD3561E526 CRC64;  
 Alignment Scores:  
 Pred. No.: 18.9 Length: 240  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.22% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-513-151-3 (1-2041) x MTR\_A\_METMA (1-240)  
 QY 1034 GCAGGATTTCAGAGCAGCTTCAA 1011  
 ||||||||||||||||||  
 Db 123 Alaaqphnginglunvalglh 130  
 RESULT 13  
 NUDC\_HAEIN STANDARD; PRT; 264 AA.  
 ID NUDC\_HAEIN  
 AC P44710;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE NADH pyrophosphatase (EC 3.6.1.-).  
 GN NUDC OR H10432  
 OS Haemophilus influenzae.  
 OC Bacteria: Proteobacteria: gamma subdivision: Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,  
 RA Weiman J.F., Phillips C.A., Spriggs T., Hebbloom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venner J.C.,  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 Rd".  
 RL Science 269:496-512(1995).  
 CC -I- CATALYTIC ACTIVITY: NADH + H(2)O = AMP + NMNH.  
 CC -I- COPACITOR: REQUIRES DIVALENT IONS: MANGANESE OR MAGNESIUM (BY  
 CC SIMILARITY).  
 CC -I- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -I- SIMILARITY: BELONGS TO THE NUDIX HYDROLASE FAMILY. NUDC SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U32726; AAC22091.1; .  
 DR TIGR: H10432; .  
 DR InterPro: IPR000086; NUDIX\_hydrolase.  
 DR Pfam: PF00293; NUDIX.1.  
 DR PRINTS: PR00302; NUDIXFAMILY.  
 DR PROSITE: PS00893; NUDIX.1.  
 KW Hydrolase; NAD; Magnesium; Manganese; Complete proteome.  
 FT DOMAIN 163 184 NUDIX BOX.  
 SQ SEQUENCE 264 AA; 30163 MW; 1720219015B9922E CRC64;  
 Alignment Scores:  
 Pred. No.: 18.6 Length: 264  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.21% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-513-151-3 (1-2041) x NUDC\_HAEIN (1-264)  
 Oy 581 GTTTGAGAAACAGATCTCT 604  
 Db 179 ValPheGluGluThreGlyIleSer 186  
 RESULT 14  
 ID PLCD\_HUMAN STANDARD: PRT; 378 AA.  
 AC Q9NR25:  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 1-acyl-sn-glycerol-3-phosphate acyltransferase delta (EC 2.3.1.51) (1-  
 DE AGP acyltransferase 4) (1-AGPAT 4) (lysophosphatidic acid  
 DE acyltransferase-delta) (LPAAT-delta) (1-acylglycerol-3-phosphate O-  
 DE acyltransferase 4).  
 GN AGPAT4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCB1\_Taxid=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Leung D.W.;  
 RT "Structure and functions of lysophosphatidic acid acyltransferases";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RN (12)  
 RC TISSUE=Spleen;  
 RA Strausberg R.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -I- FUNCTION: CONVERTS LYSPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC  
 CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION (BY  
 CC SIMILARITY).  
 CC -I- CATALYTIC ACTIVITY: Acyl-CoA + 1-acyl-sn-glycerol 3-phosphate =  
 CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.  
 CC -I- PATHWAY: De novo phospholipid biosynthesis; second step.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -I- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE  
 CC ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF156776; AAF80338.1; .  
 DR EMBL: BC020209; AAH20209.1; .  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 KW Phospholipid biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane.  
 FT TRANSMEM 11 31 POTENTIAL.  
 FT TRANSMEM 125 145 POTENTIAL.  
 FT TRANSMEM 307 327 POTENTIAL.  
 FT TRANSMEM 338 358 POTENTIAL.  
 SQ SEQUENCE 378 AA; 44021 MW; 3EFC13D196F8CDC5 CRC64;  
 Alignment Scores:  
 Pred. No.: 17.8 Length: 378  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.21% Indels: 0  
 DB: 1 Gaps: 0  
 US-09-513-151-3 (1-2041) x PLCD\_HUMAN (1-378)  
 Oy 669 AGTCTCTAACCCTGCATCCTTT 692  
 Db 334 SerSerLeuThrLeuLaseSerPhe 341  
 RESULT 15  
 ID YP28\_CAEEL STANDARD: PRT; 417 AA.  
 AC P98061;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Hypothetical zinc metalloproteinase F42A10.8 precursor (EC 3.4.24.-).  
 GN F42A10.8.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 NC NCB1\_Taxid=6239;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Latreille P.;  
 RT Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.  
 CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -I- SIMILARITY: CONTAINS 1 CUB DOMAIN.  
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sib.ch).

CC -----  
 DR EMBL: U010414; AAA19079.1; -  
 DR HSSP: P28825; 11AF.  
 DR MEROPS: M12.0PA; -  
 DR WormRep: F42A10.8; CE01299.  
 DR InterPro: IPR001506; Astacin.  
 DR InterPro: IPR000859; CUB-domain.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000130; Zn\_MTPeptide.  
 DR Pfam: PF00431; CUB; 1.  
 DR Pfam: PF01400; Astacin; 1.  
 DR SMART: SM00042; CUB; 1.  
 DR SMART: SM00235; Zmc; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS01180; CUB; 1.  
 DR PROSITE: PS00022; EGF-1; 1.  
 DR PROSITE: PS01186; EGF-2; FALSE\_NEG.  
 KW Hypothetical protein; Repeat; Hydrolase; Metalloprotease;  
 KW EGF-like domain; Zinc; Signal.  
 FT SIGNAL 1 14  
 FT CHAIN 15 417  
 FT METAL 134 134  
 FT ACT\_SITE 135 135  
 FT METAL 138 138  
 FT METAL 144 144  
 FT DOMAIN 244 274  
 FT DISULFID 284 403  
 FT DISULFID 248 259  
 FT DISULFID 251 262  
 FT DISULFID 264 273  
 FT CARBOHYD 76 76  
 FT CARBOHYD 237 237  
 FT CARBOHYD 314 314  
 SQ SEQUENCE 417 AA: 46790 MW: A0DE11E2B628C98D CRC64;  
 Alignment Scores:  
 Pred. No.: 17.6  
 Score: 8.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 1.22%  
 DB: 1  
 Length: 417  
 Matches: 8  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-09-513-151-3 (1-2041) x YP28\_CAEEL (1-417)

OY 486 CCATCCTCTTTCAGCTCACT 463  
 |||||  
 Db 305 ProSerSerSerSerSerThr 312

Search completed: April 21, 2003, 19:15:30  
 Job time : 29.5 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 19:12:25 : Search time 44 seconds

(without alignments)  
8918.653 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 661  
Sequence: 1 CTGCCATAGAGCGCGTCG.....TTTACAGAGAAAAAAA 2041

Scoring table:

OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283224 segs, 96134422 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565536

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+np.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US0515151/runal\_15042003\_141247\_27235/app\_query.fasta\_1.2183  
-DB=pir\_73 -OFMT=fasta -SUFFIX=olin2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=olin2p -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NOM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US0515151.qcgn\_1\_1\_52\_etunal\_15042003\_141247\_27235 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

1: pir\_73.\*  
2: pir\_73.\*  
3: pir\_73.\*  
4: pir\_73.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9	1.4	196	2	S47088	psad protein - Chl
2	9	1.4	382	2	S54702	twitching motility
3	9	1.4	428	2	S67176	tRNA isopentenyltr
4	9	1.4	434	2	T38664	trna isopentenyltr
5	9	1.4	698	2	G70559	probable uvrB prot
6	9	1.4	698	2	E87082	exonuclease ABC s
7	8	1.2	112	2	D81432	hypothetical prote
8	8	1.2	125	2	B53380	hypothetical prote
9	8	1.2	162	2	G75497	cell division init
10	8	1.2	213	2	B90203	hypothetical prote
11	8	1.2	213	2	B90203	ATP synthase subun
12	8	1.2	216	2	T35881	thymidine kinase (
13	8	1.2	216	2	G71512	probable cmp kinas
14	8	1.2	216	2	E81670	cytidylate kinase

15	8	1.2	246	2	F72464	hypothetical prote
16	8	1.2	263	2	G84083	ABC transporter (A
17	8	1.2	264	1	F64152	hypothetical prote
18	8	1.2	277	2	T05658	hypothetical prote
19	8	1.2	301	2	S63661	hypothetical prote
20	8	1.2	320	2	G96714	hypothetical prote
21	8	1.2	330	2	T48100	tRNA isopentenyl t
22	8	1.2	357	2	F96708	hypothetical prote
23	8	1.2	372	2	S17715	transcription acti
24	8	1.2	374	2	B82572	P-protein XF2325 (
25	8	1.2	417	2	T30863	hypothetical prote
26	8	1.2	433	2	T05693	hypothetical prote
27	8	1.2	477	2	C59098	phosphomannomutase
28	8	1.2	483	2	A18268	hypothetical prote
29	8	1.2	498	2	B70745	phosphomannomutase
30	8	1.2	507	2	S43877	trwB protein - Esc
31	8	1.2	536	2	G85488	probable transport
32	8	1.2	536	2	G90637	probable transport
33	8	1.2	580	2	S34340	NPL4 protein - yea
34	8	1.2	656	2	A64208	exonuclease ABC c
35	8	1.2	656	2	A64208	exonuclease ABC c
36	8	1.2	657	2	S73946	exonuclease ABC c
37	8	1.2	658	2	AC1761	exonuclease ABC c
38	8	1.2	658	2	A11385	exonuclease ABC c
39	8	1.2	660	2	C84099	exonuclease ABC c
40	8	1.2	661	2	G69729	exonuclease ABC c
41	8	1.2	663	2	G89848	exonuclease ABC su
42	8	1.2	665	2	A11947	exonuclease ABC c
43	8	1.2	666	2	C82932	exonuclease ABC s
44	8	1.2	669	2	S74391	exonuclease ABC c
45	8	1.2	674	2	T01309	probable serine/th

#### ALIGNMENTS

RESULT 1  
S47088  
psad protein - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 26-Aug-1999  
C:Accession: S47088  
R:Farrah, J.A.; Frank, G.; Zuber, H.; Rochaix, J.D.  
Submitted to the EMBL Data Library, June 1994  
A:Description: Cloning and sequencing of a cDNA encoding the photosystem I Psad s  
A:Reference number: S47088  
A:Accession: S47088  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-196 <FAR>  
A:Cross-references: EMBL:X79674; NID:9498823; PIDN:CAA56122.1; PID:9498824  
C:Superfamily: photosystem I chain II

Alignment Scores:  
Pred. No.: 3.84  
Score: 9.00  
Percent Similarity: 100.00%  
Best local Similarity: 100.00%  
Query Match: 1.36%  
DB: 2  
Gaps: 0

US-09-513-151-3 (1-2041) x S47088 (1-196)  
OY 130 GTTGCAGCTAGCCAGCGCTCGCGG 156  
|||||  
Db 21 ValAlaAlaArgProAlaAlaArgArg 29

RESULT 2  
S54702  
twitching motility protein Pilu PA0396 [Imported] - Pseudomonas aeruginosa (strain  
C:Species: Pseudomonas aeruginosa  
C>Date: 06-Sep-1996 #sequence\_revision 13-Mar-1997 #text\_change 24-Oct-2000  
C:Accession: S54702; B83595  
R:Whitchurch, C.B.; Mattick, J.S.

Mol. Microbiol. 13, 1079-1091, 1994  
 A:Title: Characterization of a gene, pilU, required for twitching motility but not phage  
 A:Reference number: S54702; MUID:95157252; PMID:7854122  
 A:Accession: S54702  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-382 <WH1>  
 A:Cross-references: EMBL:L27667; NID:g443685; PIDN:AAA25965.1; PID:g443686  
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: B83595  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-382 <STO>  
 A:Cross-references: GB:AE004477; GB:AE004091; NID:g9946248; PIDN:ANG03785.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: pilU; PA0396  
 C:Superfamily: twitching motility protein pilT

Alignment Scores:  
 Pred. No.: 3.52 Length: 382  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x S54702 (1-382)  
 QY 98 GGGCGCAGCGGCCGCGCAATCCAGC 124  
 Db 130 G1yAlAtHrG1yThrG1yLysSerThr 138

RESULT 3  
 S67176  
 tRNA isopentenyltransferase (EC 2.5.1.8) - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: protein O5447W; protein YOR274W  
 C:Species: *Saccharomyces cerevisiae*  
 C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
 C:Accession: S67176; A26717; S72045  
 R:Cherret, G.; Sor, F.  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S67169  
 A:Accession: S67176  
 A:Molecule type: DNA  
 A:Residues: 1-428 <CHE>  
 A:Cross-references: EMBL:Z75182; NID:g1420613; PID:e252418; PID:g1420614; MIPS:YOR274W  
 A:Experimental source: strain S288C  
 R:Naarajan, D.; Dhanich, M.E.; Martin, N.C.; Hopper, A.K.  
 Mol. Cell. Biol. 7, 185-191, 1987  
 A:Title: DNA sequence and transcript mapping of MOD5: features of the 5' region which su  
 A:Reference number: A26717; MUID:87172703; PMID:3031457  
 A:Accession: A26717  
 A:Molecule type: DNA  
 A:Residues: 1-374, R, 376-428 <NAU>  
 A:Cross-references: EMBL:M1591  
 R:Cherret, G.; Bernard, A.; Sor, F.  
 Yeast 12, 1059-1064, 1996  
 A:Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of *Saccharomyces*  
 A:Reference number: S72039; MUID:97051594; PMID:8896271  
 A:Accession: S72045  
 A>Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-428 <CHW>  
 A:Cross-references: EMBL:X89633; NID:g1279694; PIDN:CAA61780.1; PID:g1419759  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
 C:Genetics:  
 A:Gene: SGD:MOD5

A:Cross-references: MIPS:YOR274W; SGD:S0005800  
 A:Map position: 15R  
 C:Keywords: transferase

Alignment Scores:  
 Pred. No.: 3.46 Length: 428  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x S67176 (1-428)  
 QY 350 AAAATTCCTATTGTGTGGAGAGACC 376  
 Db 104 LysIleProIleValValGlyGlyThr 112

RESULT 4  
 T38664  
 tRNA isopentenyltransferase - fission yeast (*Schizosaccharomyces pombe*)  
 C:Species: *Schizosaccharomyces pombe*  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T38664  
 R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, August 1997  
 A:Reference number: Z21804  
 A:Accession: T38664  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-434 <MUR>  
 A:Cross-references: EMBL:AL109739; NID:e1534774; PIDN:CAB52278.1; GSPDB:GN00066; ;  
 A:Experimental source: strain 972h; cosmid c343  
 C:Genetics:  
 A:Gene: SPDB:SPAC343.15  
 A:Map position: 1

Alignment Scores:  
 Pred. No.: 3.46 Length: 434  
 Score: 9.00 Matches: 9  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.36% Indels: 0  
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x T38664 (1-434)  
 QY 350 AAAATTCCTATTGTGTGGAGAGACC 376  
 Db 93 LysIleProIleValValGlyGlyThr 101

RESULT 5  
 G70559  
 Probable uvrB protein - *Mycobacterium tuberculosis* (strain H37RV)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 02-Feb-2001  
 C:Accession: G70559  
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gor  
 rajandream, M.A.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holto  
 Nature 393, 537-544, 1998  
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete g  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: G70559  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-698 <COL>  
 A:Cross-references: GB:Z95554; GB:AL123456; NID:g3361771; PIDN:CAB08886.1; PID:g21  
 A:Experimental source: strain H37RV  
 C:Genetics:  
 A:Gene: uvrB  
 C:Superfamily: excinuclease ABC chain B

C:Keywords: ATP; nucleotide binding; P-loop  
F:41-48/Region: nucleotide-binding motif A (P-loop)  
F:335-340/Region: nucleotide-binding motif B  
F:339-342/Region: DEXH motif

## Alignment Scores:

Pred. No.:	3, 24	Length:	698
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x G70559 (1-698)

OY 95 CTCGGGGCCACGGCCACCGCAATCC 121

Db 40 LeuGlyAlaThrGlyThrGlyLysSer 48

## RESULT 6

E87082

exclnuclease ABC subunit B [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001

C:Accession: E87082

R:Coile, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Dutroy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Ruter, S.; Seeger, K.; Simmonds, M.; Skelton, J.; Squares, R.; Sq

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: E87082

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-698 <STO>

A:Cross-references: GB:AL450380; NID:g13093274; PIDN:CAC31768.1; GSPDB:GN00147

C:Genetics:

A:Gene: uvrB

C:Superfamily: exclnuclease ABC chain B

## Alignment Scores:

Pred. No.:	3, 24	Length:	698
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.36%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x E87082 (1-698)

OY 95 CTCGGGGCCACGGCCACCGCAATCC 121

Db 40 LeuGlyAlaThrGlyThrGlyLysSer 48

## RESULT 7

D81432

hypothetical protein Cj0148c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002

C:Accession: D81432

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chitt

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals h

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: D81432

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <PAR>

A:Cross-references: GB:AL139074; GB:AL111168; NID:g6967505; PIDN:CAB72632.1; PID:g696764

C:Genetics:

A:Gene: Cj0148c

C:Superfamily: Campylobacter jejuni hypothetical protein Cj0148c

## Alignment Scores:

Pred. No.:	46, 6	Length:	112
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.21%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x D81432 (1-112)

OY 573 GCTGCAAGTTTGAAGAAACAG 596

Db 16 AlecysLysPheLeuLysLysGln 23

## RESULT 8

C71104

hypothetical protein PH1094 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000

C:Accession: C71104

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halikawa, Y.; Hino, Y.; Yamamoto, S.

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kishida, N.

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermo

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: C71104

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-121 <KAM>

A:Cross-references: GB:AP000004; NID:g3226131; PIDN:BAA30193.1; PID:g3257510

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by c

C:Genetics:

A:Gene: PH1094

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1094

## Alignment Scores:

Pred. No.:	46, 2	Length:	121
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.22%	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3 (1-2041) x C71104 (1-121)

OY 1111 ACTCTCTGTTCTCAGCTTCATT 1088

Db 6 ThierLeuValLeuSerPheIle 13

## RESULT 9

B53380

cell division initiation protein divIC - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 20-Jun-2000

C:Accession: B53380; S60092; A69616

R:Levin, P.A.; Losick, R.

J. Bacteriol. 176, 1451-1459, 1994

A:Title: Characterization of a cell division gene from Bacillus subtilis that is

A:Reference number: A53380; MUID:94156852; PMID:8113187

A:Accession: B53380

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-125 <LRV>

A:Cross-references: GB:L23497; NID:g469178; PIDN:AAB3837.1; PID:g385177

R:Ogasawara, N.; Nakai, S.; Yoshikawa, H.

DNA Res. 1, 1-14, 1994

A:Title: Systematic sequencing of the 180 kilobase region of the Bacillus subtil

A:Reference number: S65967; MUID:96051385; PMID:7584024

A:Accession: S66092

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-125 <OGA>  
A:Cross-references: EMBL:D6185; NID:g467326; PIDN:BA05397.1; PID:g467451  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bartel  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabel, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall  
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A.; Luthers, P.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
Y, M.; Ogawa, K.; Ogihara, A.; Oudeg, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,  
A.; Schleich, S.; Schreier, R.; Scorfone, F.; Sekiguchi, J.; Sekowska, A.; Ser  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A.; Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Dancho, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A:Reference number: A69580, MUID:98044033; PMID:9384377  
A:Accession: A69616  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-125 <RUN>  
A:Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11838.1; PID:g2632329  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: divIC  
A:Start codon: TTG  
C:Superfamily: *Bacillus subtilis* cell division initiation protein divIC

Alignment Scores:  
Pred. No.: 46 Length: 125  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.21% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x B53380 (1-125)  
QY 1193 AACCAACTGAGAGAGAGAGACA 1216  
|||||  
DB 24 AaNgInLeuLysLysArGArGArG 31

RESULT 10  
C75497  
Hypothetical protein - *Deinococcus radiodurans* (strain R1)  
C:Species: *Deinococcus radiodurans*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: C75497  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
A:Reference number: A75250, MUID:20036896; PMID:10567266  
A:Accession: C75497  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-162 <WHI>  
A:Cross-references: GB:AE001919; GB:AE000513; NID:g6458307; PIDN:AAF10196.1; PID:g645831  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0609  
A:Map position: 1

Alignment Scores:  
Pred. No.: 44.4 Length: 162  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.21% Indels: 0

DB: 2 Gaps: 0  
US-09-513-151-3 (1-2041) x C75497 (1-162)  
QY 107 GGCCAGCGCAATTCAGCGCGC 130  
|||||  
DB 84 GlyThrGlyLysSerThrLeuAla 91

RESULT 11  
B90203  
ATP synthase subunit D (atpD) [imported] - *Sulfolobus solfataricus*  
C:Species: *Sulfolobus solfataricus*  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 02-Aug-2002  
C:Accession: B90203  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweez, M.J.  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Re  
arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: *Sulfolobus solfataricus* complete genome.  
A:Reference number: A99139  
A:Accession: B90203  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-213 <KUR>  
A:Cross-references: GB:AE006641; NID:g13813731; PIDN:AAK40881.1; GSPDB:GN00155  
C:Genetics:  
A:Gene: atpD  
C:Superfamily: H<sup>+</sup>-transporting ATPase chain D

Alignment Scores:  
Pred. No.: 42.8 Length: 213  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.21% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x B90203 (1-213)  
QY 1858 ATCTACAGAGAGAGAGATCA 1881  
|||||  
DB 203 IleLeuGlnArgArgArgGluSer 210

RESULT 12  
T35881  
thymidine kinase (EC 2.7.1.21) [similarity] - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jun-2002  
C:Accession: T35881  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Raftery, M.A.  
submitted to the EMBL Data Library, November 1997  
A:Reference number: 221592  
A:Accession: T35881  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-216 <OL>  
A:Cross-references: EMBL:AL009204; PIDN:CAA15802.1; GSPDB:GN00070; SCOEDB:SC9B10.1  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC9B10.12  
C:Superfamily: *Escherichia coli* thymidine kinase  
C:Keywords: phosphotransferase

Alignment Scores:  
Pred. No.: 42.7 Length: 216  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.21% Indels: 0  
DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x T35881 (1-216)



OY 113 GGCAATCCAGCTGGCTGCAG 136  
 |||||||||||||||||||||  
 Db 14 GlyThrcGlyLysSerThrLeuAla 21

## RESULT 13

G71512  
 Probable cmp kinase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
 C:Accession: G71512  
 R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
 A:Reference number: A71512; MUID:99008089; PMID:9784136  
 A:Accession: G71512  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-216 <ARN>  
 A:Cross-references: GB:AE001319; GB:AE001273; NID:93328881; PIDN:ACG68052.1; PID:9332888  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: cmk  
 C:Superfamily: cytidylate kinase cmk

Alignment Scores:  
 Pred. No.: 42.7 Length: 216  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.21% Indels: 0  
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x G71512 (1-216)

OY 107 GGCACCGCAATCCAGCTGGCG 130  
 |||||||||||||||||||||  
 Db 10 GlyThrcGlyLysSerThrLeuAla 17

## RESULT 14

E81670  
 cytidylate kinase TC0737 [imported] - Chlamydia muridarum (strain Nigg)  
 C:Species: Chlamydia muridarum, Chlamydia trachomatis MOPn  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-May-2000  
 C:Accession: E81670  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39.  
 A:Reference number: A81500; MUID:20150255; PMID:10684935  
 A:Accession: E81670  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-216 <RET>  
 A:Cross-references: GB:AE002342; GB:AE002160; NID:97190763; PIDN:AAF39547.1; PID:9719076  
 A:Experimental source: strain Nigg (MOPn)  
 C:Genetics:  
 A:Gene: TC0737  
 C:Superfamily: cytidylate kinase cmk

Alignment Scores:  
 Pred. No.: 42.7 Length: 216  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.21% Indels: 0  
 DB: 2 Caps: 0

US-09-513-151-3 (1-2041) x E81670 (1-216)

OY 107 GGCACCGCAATCCAGCTGGCG 130  
 |||||||||||||||||||||  
 Db 10 GlyThrcGlyLysSerThrLeuAla 17

## RESULT 15

F72464  
 hypothetical protein APE2361 - Aeropyrum pernix (strain KI)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C:Accession: F72464  
 R:Kakarbayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Iwa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon,  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: F72464  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-246 <KAN>  
 A:Cross-references: DDBJ:AP000064; NID:95105945; PIDN:BAAB1374.1; PID:95106063  
 A:Experimental source: strain KI  
 C:Genetics:  
 A:Gene: APE2361  
 C:Superfamily: peroxidase

Alignment Scores:  
 Pred. No.: 42 Length: 246  
 Score: 8.00 Matches: 8  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.21% Indels: 0  
 DB: 2 Gaps: 0

US-09-513-151-3 (1-2041) x F72464 (1-246)

OY 752 CTTGCTGCTGGCTCTTGAGAGCA 775  
 |||||||||||||||||||||  
 Db 193 LeuAlaAlaGlyLeuLeuGluGlu 200

Search completed: April 21, 2003, 19:21:02  
 Job time : 51 secs



Seq primer: M13 FORWARD  
POLY-A-Tes.

FEATURES  
source

Location/Qualifiers  
1..624  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="5845067"  
/clone\_11b="NCI-GAP\_E11"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; NCI-GAP\_E11 is a normalized cDNA library containing the following tissue(s): Chondrosarcoma. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is  
ACACTGCAC.  
TAG\_L1B-U1-H-E11  
TAG\_TISSUE=chondrosarcoma  
TAG\_SEO=ACACTGCAC"

BASE COUNT 203 a 128 c 97 g 195 t 1 others

ORIGIN

Query Match 29.7%; Score 606.8; DB 14; Length 624;  
Best Local Similarity 99.4%; Pred. No. 2e-111;  
Matches 619; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1419 AGGAGTCATGCGGGAATCTCTGCATACGAGAAAGCTCCGCCATTTCTTTGATGT 1478  
DB 624 AGGAGTCATGCGGGAATCTCTGCATACGAGAAAGCTCCGCCATTTCTTTGATGT 565  
QY 1479 GGTTTAAAGTCTCAGCTTCTCTAATATAGAAACAGCAGTCTTGACAGCTCTGTG 1538  
DB 564 GGTTTAAAGTCTCAGCTTCTCTAATATAGAAACAGCAGTCTTGACAGCTCTGTG 505  
QY 1539 GCTGATGTCCTGGAATATGATGATGTCAGGAAAGCATTCTTTCTTTGACCTTAA 1598  
DB 504 GCTGATGTCCTGGAATATGATGATGTCAGGAAAGCATTCTTTCTTTGACCTTAA 445  
QY 1599 AGGTCATATTAAGAGAGACAGATCCACATTTTATATCATAGATCTCTCTTGT 1658  
DB 444 AGGTCATATTAAGAGAGACAGATCCACATTTTATATCATAGATCTCTCTTGT 385  
QY 1659 GGTGAATACAGGATGATGATGATCCCTTAAAGAGTTTATGTCCTGACCTGGCT 1718  
DB 384 GGTGAATACAGGATGATGATGATCCCTTAAAGAGTTTATGTCCTGACCTGGCT 326  
QY 1719 AAAATATCTAATTTCCAGATGCTTTTGTAGATGATGAGATTTTGTAGCCACATAT 1778  
DB 325 AAAATATCTAATTTCCAGATGCTTTTGTAGATGATGAGATTTTGTAGCCACATAT 266  
QY 1779 GGGAGTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATTGAGATGTTAAGTG 1838  
DB 265 GGGAGTCTAGATTTGAGTGAATGGCAGAAAGGCCATCTCCATTGAGATGTTAAGTG 206  
QY 1839 AACCAAACTAGTTCTCGAATTTCTACAGAGAGAGGAGTACAGATGAGGAAAGCTGTA 1898  
DB 205 AACCAAACTAGTTCTCGAATTTCTACAGAGAGAGGAGTACAGATGAGGAAAGCTGTA 146  
QY 1899 CATAGAGCTTGAAGACCAAGACTTTGAATTTGCGAGCTGCATGCTGATTTAT 1958  
DB 145 CATAGAGCTTGAAGACCAAGACTTTGAATTTGCGAGCTGCATGCTGATTTAT 86  
QY 1959 CACTGCTGCTTTCTATTGATGATCAATCTAATTTTATGAGACTTTAAATAAGAA 2018

DB 85 CACTGCTGCTTTCTATTGATGATCAATCTAATTTTATGAGACTTTAAATAAGAA 26  
QY 2019 AAATTACAGAAAAA 2041  
DB 25 AAATTACAAAAA 3

RESULT 14  
AV759288  
LOCUS AV759288 639 bp mRNA linear EST 19-OCT-2000  
DEFINITION MDS Homo sapiens cDNA clone MDSBLH01.5, mRNA sequence.  
ACCESSION AV759288  
VERSION AV759288.1 GI:10917136  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Yang,Y., Gao,G., Zhang,Q., Chen,S., Han,Z. and Chen,Z.  
TITLE Homo sapiens cDNA MDS clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES  
source

1..639  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="MDSBLH01"  
/clone\_11b="MDS"  
/tissue\_type="bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/note="Vector: pT73-Pac; Site\_1: SfiI; Site\_2: SfiI"

BASE COUNT 166 a 158 c 175 g 139 t 1 others

ORIGIN

Query Match 29.4%; Score 599.8; DB 10; Length 639;  
Best Local Similarity 98.4%; Pred. No. 5.1e-110;  
Matches 628; Conservative 0; Mismatches 3; Indels 7; Gaps 2;

QY 1 CTGCGATTAAGATGGCGTCCGTCGTCGAGAGAGATTCCTGTGGAGTGGGCTCAG 60  
DB 5 CTGCGATTAAGATGGCGTCCGTCGTCGAGAGAGATTCCTGTGGAGTGGGCTCAG 64  
QY 61 GGGCTGCAACGAGCCCTACCTCTTGTAGTATCTCGGGGCCACGGGCAACCGCAATC 120  
DB 65 GGGCTGCAACGAGCCCTACCTCTTGTAGTATCTCGGGGCCACGGGCAACCGCAATC 124  
QY 121 CACGCTGGGCTTGACAGTGGCCAGCGGCTCGCGGTGAGATGTGACGCTACTTCAT 180  
DB 125 CACGCTGGGCTTGACAGTGGCCAGCGGCTCGCGGTGAGATGTGACGCTACTTCAT 184  
QY 181 GGAAGTCTAAGAGGCTGACATCATCAACCAAGGTTCTGCCAAGAGAGAGAAAT 240  
DB 185 GGAAGTCTAAGAGGCTGACATCATCAACCAAGGTTCTGCCAAGAGAGAGAAAT 244  
QY 241 CTGCGGCAACCATGATGATGATTTGTGATCTCTTGTGACCAATTACAGATGCTGA 300  
DB 245 CTGCGGCAACCATGATGATGATTTGTGATCTCTTGTGACCAATTACAGATGCTGA 304  
QY 301 CTTCAGAAATAGAGCACTGCTCTGATTAAGATATATTTGGCCGAGACAAAATCTTAT 360  
DB 305 CTTCAGAAATAGAGCACTGCTCTGATTAAGATATATTTGGCCGAGACAAAATCTTAT 364

Db 482 TTGTACTGCACAAACGCTTAAGCCAGGTGGAGCCAGAACATGGCTGCCAATGTCATCA 541

QY 548 CAGCAAAAGCGAAAGTGGCGAGCTTGCAGTCTTGAAGAACAGGAATCTGCAT 607

Db 542 CATGAAGAAAGCGAAAGTGGCGAGCTTGCAGTCTTGAAGAACAGGAATCTGCAT 601

QY 608 AGTGAATTTCTCCATGCTCAACATACGGAAGAGTGGTGGTCCCTTGGAGGCTCTG 667

Db 602 AGTGAATTTCTCCATGCTCAACATACGGAAGAGTGGTGGTCCCTTGGAGGCTCTG 661

QY 668 AAGTCTCTAACCCTTGCATCCCTTGGCTTCAATGACGAGCAGTCTAGATGAGCG 727

Db 662 AGTCTCTAACCCTTGCATCCCTTGGCTTCAATGACGAGCAGTCTAGATGAGCG 718

QY 728 TTGGATTAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 763

Db 719 TGGCATAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATG 754

RESULT 12

LOCUS BC034660 772 bp mRNA linear EST 24-JAN-2001

DEFINITION 602300022F1 NIH\_MGC.87 Homo sapiens cDNA clone IMAGE:4394260 5', mRNA sequence.

ACCESSION BC034660

VERSION BC034660.1 GI:12428227

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 772)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DRP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
Plate: LLM10089 row: 1 column: 05  
High quality sequence stop: 692.

FEATURES

source

location/Qualifiers

1..772

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="4394260"

/clone\_lib="NIH\_MGC.87"

/tissue\_type="mammary adenocarcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: breast; Vector: pCMV-SPORT6; Site: 1; NotI; Site: 2; SalI; Cloned unidirectionally; oligo-dt primed. Average insert size 1.383 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT 224 a 155 c 178 g 214 t 1 others

ORIGIN

Query Match 30.88; Score 628; DB 12; Length 772;

Best Local Similarity 95.18; Pred. No. 1.le-115;

Matches 735; Conservative 0; Mismatches 26; Indels 12; Gaps 8;

QY 1001 GAGCTGCTGCTGAACCTGCTGGAATGCTGCAAGTTTCATCCAGGGCCACAAGCT 1060

Db 2 GAGCTGCTGCTGAACCTGCTGGAATGCTGCAAGTTTCATCCAGGGCCACAAGCT 60

QY 1061 ACAGCCACCTCCATTAAGATGCCATACATGAGAGCTGAGAACAGAGATTAACCTG 1120

Db 61 ACAGCCACCTCCATTAAGATGCCATACATGAGAGCTGAGAACAGAGATTAACCTG 120

QY 1121 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGGAATGGCGACGACATMAATCC 1180

Db 121 TGTGACCTCTGTGATGCAATCATCATTTGGGATCGGAATGGCGACGACATMAATCC 180

QY 1181 AAATCCCACTTGAACCACTGAAGAAAAGAAAGATGGAGTGCATGCTGTCAACACC 1240

Db 181 AAATCCCACTTGAACCACTGAAGAAAAGAAAGATGGAGTGCATGCTGTCAACACC 240

QY 1241 ATGAAGATGAGAGTGTTCCTCCAGACTATAACAAAGAACCTTAAGAGGAGATCCCA 1300

Db 241 ATGAAGATGAGAGTGTTCCTCCAGACTATAACAAAGAACCTTAAGAGGAGATCCCA 300

QY 1301 GGCAGATATGATCAAGAGCTGGAATGAGCGGTTTAAGAGACATGTCAGTGGCTTTGGA 1360

Db 301 -GGCAGATATGATCAAGAGCTGGAATGAGCGGTTTAAGAGACATGTCAGTGGCTTTGGA 359

QY 1361 AAGTGTGGGGATCCAGTTCAGAGAGGAGGGATGTTGTCTCCAGTCTGGGCAAG 1420

Db 360 AAGTGTGGGGATCCAGTTCAGAGAGGAGGGATGTTGTCTCCAGTCTGGGCAAG 419

QY 1421 GAGTGTATGCGGATTCCTGCTGATACGACAGAAAAGCTCCACATTTCTTTGATGTG 1480

Db 420 GAGTGTATGCGGATTCCTGCTGATACGACAGAAAAGCTCCACATTTCTTTGATGTG 479

QY 1481 TTTTAAGTCTCAAGCTCTCTATATATGAAGAACAGAGCTTGTGACGCTCTGTGTGG 1540

Db 480 TTTTAAGTCTCAAGCTCTCTATATATGAAGAACAGAGCTTGTGACGCTCTGTGTGG 539

QY 1541 TGATGTCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599

Db 540 TGATGTCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 599

QY 1600 GGTCTCTATATTAAGACA -GCACAGTTCACCA -TTTTATATCATAGATGCTCTG 1657

Db 600 GGTCTCTATATTAAGACA -GCACAGTTCACCA -TTTTATATCATAGATGCTCTG 659

QY 1658 TGGTGAATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1717

Db 660 TGGTGAATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 715

QY 1718 TAAAT 1770

Db 716 TAAAT 765

RESULT 13

LOCUS B0003256/c 624 bp mRNA linear EST 26-MAR-2002

DEFINITION UI-H-E11-ayx-n-12-0-UI.s1 NCI\_CGAP\_E11 Homo sapiens cDNA clone. IMAGE:5845067 3', mRNA sequence.

ACCESSION B0003256

VERSION B0003256.1 GI:19728156

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 624)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone distribution: Information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
The following repetitive elements were found in this cDNA sequence: 1-36, >POLY\_A#simple\_repeat (matched complement)



451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu

Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

## FEATURES

## Source

Location/Qualifiers

1. 724

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="UT-E-EO1-ab-b-20-0-UT"

/clone\_lib="UT-E-EO1"

/tissue\_type="fetal eye"

/dev\_stage="fetal"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site\_1: EcoR I; Site\_2: Not I; UT-E-EO1 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGGTAATCC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 208 a 132 c 175 g 208 t 1 others  
ORIGIN

## Query Match

Best Local Similarity 33.28; Score 677.8; DB 14; Length 724;  
Matches 715; Conservative 0; Mismatches 8; Indels 3; Gaps 3;

QY 1117 CCTGTGACCTGTGTGATCATCATCTGGGATCGCAATGGCGAGCCACATATA 1176  
Db 1 CCTGTGACCTGTGTGATCATCATCTGGGATCGCAATGGCGAGCCACATATA 60  
QY 1177 ATCCAAATCCCACTGGAACCACTGAAGAAAGAAAGATTGGACTCAGATGCTCAA 1236  
Db 61 ATCCAAATCCCACTGGAACCACTGAAGAAAGAAAGATTGGACTCAGATGCTCAA 120  
QY 1237 CACCATGAAAGTCAAGAGTGTTCCTCCAGACTATACAAAGACCTAAAGGAGGATC 1296  
Db 121 CACCATGAAAGTCAAGAGTGTTCCTCCAGACTATACAAAGACCTAAAGGAGGATC 180  
QY 1297 CCCAGGCGAGATGATCAAGAGTGAAGTGAAGGAGGATTAAGAGATGTCAGTGGCTT 1356  
Db 181 CCCAGGCGAGATGATCAAGAGTGAAGTGAAGGAGGATTAAGAGATGTCAGTGGCTT 240  
QY 1357 TGAAGAGTGTGGGATCCAGTTCAGAGAGGAGGAGTATGTTGCTCCAGTCTGGGC 1416  
Db 241 TGAAGAGTGTGGGATCCAGTTCAGAGAGGAGGAGTATGTTGCTCCAGTCTGGGC 300  
QY 1417 AAAGAGTGTGATGCGGAATTCCTGATAGCAGAAAAGCTCCACCACTTTTCTTTTAT 1476  
Db 301 AAAGAGTGTGATGCGGAATTCCTGATAGCAGAAAAGCTCCACCACTTTTCTTTTAT 360  
QY 1477 GTGTTTAAAGTCTCAGTCTCTATATAGAAACAGCAGCTGTGACGCTCTTG 1536  
Db 361 GTGTTTAAAGTCTCAGTCTCTATATAGAAACAGCAGCTGTGACGCTCTTG 420  
QY 1537 TGCGTGAATGTGCTGGAATGATGATGTCAGAAAGCA-TTTTTTTTTTCTTGAACCT 1595  
|||||

Db 421 TGCGTGAATGTGCTGGAATGATGATGTCAGAAAGCAATTTTTTTTCTTGAACCT 480  
QY 1596 TAAAGTCTCTATTTTAAACAGCAGATTCACATTTTATACATGAGATCTCTT 1655  
Db 481 TAAAGTCTCTATTTTAAACAGCAGATTCACATTTTATACATGAGATCTCTT 540  
QY 1656 TGTGTGAATCCAGAGATTCAGTCCATCCCTTTAAAGATTTTATGTCGACCTG 1715  
Db 541 TGTGTGAATCCAGAGATTCAGTCCATCCCTTTAAAGATTTTATGTCGACCTG 599  
QY 1716 GCTAAATATATATATTCAGATCTTTTGTAGATGATGATGATGATGAT 1775  
Db 600 GCTAAATATATATATTCAGATCTTTTGTAGATGATGATGATGATGATGAT 659  
QY 1776 ATTTGAGATTCATATTTGATGATGATGATGATGATGATGATGATGAT 1835  
Db 660 ATTTGAGATTCATATTTGATGATGATGATGATGATGATGATGATGAT 718  
QY 1836 GTGAAC 1841  
Db 719 GTGAAC 724

RESULT 10  
Bg612651  
LOCUS  
DEFINITION  
602640078F1 NIH\_MGC\_61 Homo sapiens cDNA clone IMAGE:4771338 5',  
mRNA sequence.  
ACCESSION  
Bg612651  
VERSION  
Bg612651.1 GI:13664022  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 869)  
AUTHORS  
NIH-MGC http://mgc.nci.nih.gov/  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
CONTACT  
Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM1639 row: k column: 19  
High quality sequence stop: 684.

## FEATURES

## Source

Location/Qualifiers

1. 869

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:4771338"

/clone\_lib="NIH\_MGC\_61"

/tissue\_type="embryonal carcinoma"

/lab\_host="DH10B (T1 phage-resistant)"

/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site\_1: SfiI (ggccgcctggcc); Site\_2: SfiI (ggccatattggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGCGGAGCGGCGGCGCATG-dt(30)-BN-3 (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT 259 a 184 c 206 g 220 t  
ORIGIN

```

|||||
Db 240 GGATCCCTGTGTGACCAATTCACAGTGTGAGCTTGAGAAATGACCACTGCTGAT 299
QY 328 TGAAGATATATTTGGCCGAGACAAATTCCTATTTGTTGGAGGAGCAATTTTACAT 387
Db 300 TGAAGATATATTTGGCCGAGACAAATTCCTATTTGTTGGAGGAGCAATTTTACAT 359
QY 388 TGAATCTCTCTGTGAAAGTTCTTGTCAATACCAGCCCCAGAGATGGGCACTGAGAA 447
Db 360 TGAATCTCTCTGTGAAAGTTCTTGTCAATACCAGCCCCAGAGATGGGCACTGAGAA 419
QY 448 AGGATTTGACCGAAAGTGGAGCTTGAAGAGAGATGGTCTTACTTACAAAGCCCT 507
Db 420 AGGATTTGACCGAAAGTGGAGCTTGAAGAGAGATGGTCTTACTTACAAAGCCCT 479
QY 508 AACCCAGGTGGACCCAGAAATGGTGCCAGCTGCATCCATGACAAAGCAAGTGGC 567
Db 480 AACCCAGGTGGACCCAGAAATGGTGCCAGCTGCATCCATGACAAAGCAAGTGGC 539
QY 568 CAGAGCTTGCAGATTTTGAAGAAACAGAAATCTCTCAATGTAATTTTCCATGCTCA 627
Db 540 CAGAGCTTGCAGATTTTGAAGAAACAGAAATCTCTCAATGTAATTTTCCATGCTCA 599
QY 628 ACATACGGAAGAAAGTGGTGGTCCCTTGAAGTCTCTCAAGTTCTTAACCTTGCAT 687
Db 600 ACATACGGAAGAAAGTGGTGGTCCCTTGAAGTCTCTCAAGTTCTTAACCTTGCAT 659
QY 688 CCTTGGCTTCATGCTGACGAGCAGTTCTAGATGAGCGCTTGATTAAGAGGTGATGA 747
Db 660 CCTTGGCTTCATGCTGACGAGCAGTTCTAGATGAGCGCTTGATTAAGAGGTGATGA 719
QY 748 CATGC 752
Db 720 CTGTC 724

RESULT 8
LOCUS A1133396 725 bp mRNA linear EST 11-NOV-1999
DEFINITION HA2003 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
ACCESSION A1133396
VERSION A1133396.1 GI:6360712
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 725)
AUTHORS Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
COMMENT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yyt48@yahoo.com

FEATURES
Location/Qualifiers
1..725
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="Vector: pCDNA1"

BASE COUNT 220 a 177 c 124 g 204 t
ORIGIN
Query Match 33.7%; Score 688.2; DB 9; Length 725;

```

```

Best Local Similarity 99.3%; Pred. No. 1e-127;
Matches 712; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 1274 AAGAACCTTAAGAGAGAGGAT-CCCGAGGCGAAGATCAAGAGCTGAATACACGCT 1332
Db 717 AAGAACTTAAGAGAGAGGATCCCGAGGCGAAGATCAAGAGCTGAATACACGCT 658
QY 1333 TTAAGACATATGTCAGTGGCTTTTGAAGAGTGGGGATCCAGTTCAGAGGAGGG 1392
Db 657 TTAAGACATATGTCAGTGGCTTTTGAAGAGTGGGGATCCAGTTCAGAGGAGGG 598
QY 1393 GTATGTTTGTCTCCAGTCTGGGCAAGAGAGTCTATGCGGAATTTCTGATACGAGA 1452
Db 597 GTATGTTTGTCTCCAGTCTGGGCAAGAGAGTCTATGCGGAATTTCTGATACGAGA 538
QY 1453 AAGCTCCACCATTTTCTTTGATGATGTTTAAAGTCTACGTTCTTAATTAAGAAC 1512
Db 537 AAGCTCCACCATTTTCTTTGATGATGTTTAAAGTCTACGTTCTTAATTAAGAAC 478
QY 1513 AGCAGTCTTGTACGCTCTGTGTGCTGATGCTGTAATGATGATGTCAGAGAA 1572
Db 477 AGCAGTCTTGTACGCTCTGTGTGCTGATGCTGTAATGATGATGTCAGAGAA 418
QY 1573 GCATTTTCTTTCTTTGAACCTTAAGAGTCTATTTAAGACAGACAGATTCACA 1632
Db 417 GCATTTTCTTTCTTTGAACCTTAAGAGTCTATTTAAGACAGACAGATTCACA 358
QY 1633 TTTTATACATGAGAGATCTTGTGTGATATCCAGAGATGAGTGCATCCCTTAAA 1692
Db 357 TTTTATACATGAGAGATCTTGTGTGATATCCAGAGATGAGTGCATCCCTTAAA 299
QY 1693 GAAGTTTATGTCCTGACCTGCTGCTAAATTTATCTAATTTCCAGATGCTTTGTAGATG 1752
Db 298 GAAGTTTATGTCCTGACCTGCTGCTAATTTATCTAATTTCCAGATGCTTTGTAGATG 239
QY 1753 ACTGAATATTTGTGAGCCACATATTTGGAGCTCTAGATTTGAGTAAATGGCAGAAAG 1812
Db 238 ACTGAATATTTGTGAGCCACATATTTGGAGCTCTAGATTTGAGTAAATGGCAGAAAG 179
QY 1813 GCCATCTCCATGAGATGATTAAGTAAACCAATAGTCTCCGAAATTTCTACAGAGAG 1872
Db 178 GCCATCTCCATGAGATGATTAAGTAAACCAATAGTCTCCGAAATTTCTACAGAGAG 119
QY 1873 AGGGAATCAGATGAGAGAGCTGTGACATAGAGCTTGAAGACCAAGCTTTGAATTTG 1932
Db 118 AGGGAATCAGATGAGAGAGCTGTGACATAGAGCTTGAAGACCAAGCTTTGAATTTG 59
QY 1933 CGAGCTGCTATGTTGAGTATTTATACAGTCTGCTTTTCTATTTGAGTTACAAATCT 1989
Db 58 CGAGCTGCTATGTTGAGTATTTATACAGTCTGCTTTTCTATTTGAGTTACAAATCT 2

RESULT 9
LOCUS BM721352 724 bp mRNA linear EST 01-MAR-2002
DEFINITION UI-E-EO1-a1b-20-0-UI.r1 UI-E-EO1 Homo sapiens cDNA clone
ACCESSION BM721352
VERSION BM721352.1 GI:19041207
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 724)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa

```

```

/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/Note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
U1-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTCGTCAGGT.
TAG_L1B-U1-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTGCTCAGGT"
BASE COUNT      231 a      172 c      116 g      224 t
ORIGIN
Query Match      35.6%; Score 727.4; DB 14; Length 743;
Best Local Similarity 99.7%; Pred. No. 1.4e-135;
Matches 739; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1301 GGGCAGATGATCAAGAGCTGAAGTCAAGCCTTAAAGACATGTCAGTGGCCTTTGGA 1360
DB 743 GGGCAGATGATCAAGAGCTGAAGTCAAGCCTTAAAGACATGTCAGTGGCCTTTGGA 684
QY 1361 AAGGTGGTGGGATCCAGTTCAGAGAGGAGGATGTTGTCCTCCAGTCTGGGCAAG 1420
DB 683 AAGGTGGTGGGATCCAGTTCAGAGAGGAGGATGTTGTCCTCCAGTCTGGGCAAG 624
QY 1421 GAGTGTATGCGGAATTCCTGATAGCAGAAAAGCTCCACCAATTTCTTTGATGTGG 1480
DB 623 GAGTGTATGCGGAATTCCTGATAGCAGAAAAGCTCCACCAATTTCTTTGATGTGG 564
QY 1481 TTTTAAAGTCTCAGCTCTCTAATAATGAACACAGAGTCTTTCAGTCTCTGTGGC 1540
DB 563 TTTTAAAGTCTCAGCTCTCTAATAATGAACACAGAGTCTTTCAGTCTCTGTGGC 504
QY 1541 TGATGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1600
DB 503 TGATGTCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 444
QY 1601 GTTCTATTTTAAAGACAGACAGATTCACATTTTATACATGAGATCTTCTTTGG 1660
DB 443 GTTCTATTTTAAAGACAGACAGATTCACATTTTATACATGAGATCTTCTTTGG 384
QY 1661 TGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1720
DB 383 TGAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 325
QY 1721 AATTATCTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATG 1780
DB 324 AATTATCTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATG 265
QY 1781 GAGTTCTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1840
DB 264 GAGTTCTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 205
QY 1841 CCAAACTAGATTTCTGGAATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
DB 204 CCAAACTAGATTTCTGGAATTTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 145
QY 1901 TAGACTTGAAGACCAAGAGATTTGAATTTGAGAGCTGATGATGATGATGATGATGAT 1960
DB 144 TAGACTTGAAGACCAAGAGATTTGAATTTGAGAGCTGATGATGATGATGATGATGAT 85
QY 1961 CTCGTCCTTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2020

```

```

DB 84 CTCGTCCTTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 25
QY 2021 ATTACAGAGAAAAA 2041
DB 24 ATTACAGAGAAAAA 4
RESULT 7
LOCUS BE315223
DEFINITION BE315223 725 bp mRNA linear EST 26-OCT-2000
ACCESSION 601141778F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3141643 5'
VERSION BE315223
KEYWORDS BE315223.1 GI:9145717
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 725)
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs.femail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at: image.lmnl.gov
Plate: LMCL11 row: k column: 20
High quality sequence stop: 720.
FEATURES
source
1..725
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3141643"
/clone_1b="NIH_MGC_9"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: ovary; Vector: pOT8; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      186 a      175 c      197 g      167 t
ORIGIN
Query Match      33.8%; Score 690.6; DB 10; Length 725;
Best Local Similarity 97.9%; Pred. No. 3.3e-128;
Matches 710; Conservative 0; Mismatches 14; Indels 1; Gaps 1;
QY 28 TGACGAGACATTCCTGAGGAGTGGGCTGAGGGGCTGACAGGAGCCCTACCTTGT 87
DB 1 TGACGAGACATTCCTGAGGAGTGGGCTGAGGGGCTGACAGGAGCCCTACCTTGT 60
QY 88 AGTATTCCTGAGGAGTGGGCTGAGGAGTGGGCTGAGGAGTGGGCTGAGGAGTGGGCTG 147
DB 61 AGTATTCCTGAGGAGTGGGCTGAGGAGTGGGCTGAGGAGTGGGCTGAGGAGTGGGCTG 120
QY 148 GCTGGGGGAGATGCTGACGCTGATCCTCATGAGGCTATGAGAGGCTAGACATCAT 207
DB 121 GCTGGGGGAGATGCTGACGCTGATCCTCATGAGGCTATGAGAGGCTAGACATCAT 180
QY 208 CACCAACAGTTTCTGCCCAAGACAGAGATGTCGGGAGCAGCAGTGTGATGATGATGAT 267
DB 181 CACCAACAGTTTCTGCCCAAGACAGAGATGTCGGGAGCAGCAGTGTGATGATGATGAT 239
QY 268 GGATCCTCTTGTGACCAATTACAGAGTGTGATGATGATGATGATGATGATGATGATGAT 327

```



LOCUS	DEFINITION	AGENCY	VERSION	KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
BM800217	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
LOCUS	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
DEFINITION	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
AGENCY	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
VERSION	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
KEYWORDS	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
ORGANISM	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
REFERENCE	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
AUTHORS	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
TITLE	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
JOURNAL	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
COMMENT	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
FEATURES	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
SOURCE	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
BASE COUNT	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
ORIGIN	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Query Match	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Best Local Similarity	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Matches	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Conservative	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Mismatches	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Indels	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Gaps	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Score	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
DB	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Length	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Score	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
DB	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Length	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Score	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
DB	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Length	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Score	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
DB	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Length	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Score	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
DB	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Length	BM800217	AGENCOURT	6416157	NIH_MGC_71	Homo sapiens	CDNA clone	IMAGE:5531248				
Score	BM800217	AGENCOURT	6416157	NIH_MGC_71							

QY	1339	GACATGTCACAGTGGGCTTTGGAAAGTGTGTGGGATCCAGTTCACAGAGGAGGATATGT	1398
Db	501	GACATGTCACAGTGGGCTTTGGAAAGTGTGTGGGATCCAGTTCACAGAGGAGGATATGT	560
QY	1399	TTGTCTCCCACTGTGGGCAAGAGAGTCTATGCGGAAATTCCTGTGCATAGCAGAAAAAGCTC	1458
Db	561	TTGTCTCCCACTGTGGGCAAGAGAGTCTATGCGGAAATTCCTGTGCATAGCAGAAAAAGCTC	620
QY	1459	CCACCATTTTCTTTTGATGTGGTTTAAAGTCTCACGTTCTCTATATATGAAGAACGACAGG	1518
Db	621	CCACCATTTTCTTTTGATGTGGTTTAAAGTCTCACGTTCTCTATATATGAAGAACGACAGG	680
QY	1519	TCTGTGTCAGCCCTGTGTGGCTGTATGTGTGTGCAAAATGATGTACTTACAGAAACATTT	1578
Db	681	TCTGTGTCAGCCCTGTGTGGCTGTATGTGTGTGCAAAATGATGTACTTACAGAAACATTT	740
QY	1579	TTTTTTCTTTGAACCTTAAAGGTTCTATATATTAAGACAGACAGATTCACATTTTAA	1638
Db	741	TTTTTTCTTTGAACCTTAAAGGTTCTATATATTAAGACAGACAGATTCACATTTTAA	798
QY	1639	TACATGAGAGATCTTTTGTGTGTATACACAGAGATTGACTGCATCCCTTTAAAGAAAGTT	1698
Db	799	TACATGAGAGATCTTTTGTGTGTATACACAGAGATTGACTGCATCCCTTTAAAGAAAGTT	858
QY	1699	TTATGT-CCCTGACTCT-GGCTAAATATATCAATTTTCCA	1736
Db	859	TTATGTCCCTGACTCTGGCTAAATATATCAATTTTCCA	898

RESULT 6  
BM975520/c

LOCUS  
DEFINITION  
UT-CE-EN1-acw-c-07-0-UI.s1 UT-CE-EN1 Homo sapiens cDNA clone  
UT-CE-EN1-acw-c-07-0-UI 3, mRNA sequence.  
BM975520  
BM975520.1 GI:19593111  
EST.

KEYWORDS  
SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 743)  
Ronald M.F., Lennon, G., and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
9704447  
Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul.mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-56, >AT-rich<low\_complexity (matched complement)  
Seq primer: M13 FORWARD  
POLYA-yes.

FEATURES  
SOURCE  
Location/Qualifiers  
1..743  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UT-CE-EN1-acw-c-07-0-UI"  
/clone\_lib="UT-CE-EN1"  
/tissue\_type="Primary Lung Cystic Fibrosis Epithelial  
Cells"

						/lab_host="DH10B"
						/note="Organ: brain; Vector: PCMV-SPORE; Site.1: NotI;
						Site.2: EcoRV (destroyed); RNA source anonymous pool of 6
						male brains, age range 23-27 yo. Library is oligo-dT
						primed and directionally cloned (EcoRV site is destroyed
						upon cloning). Average insert size 1.5 kb. Insert size
						range 1-3 kb. Library is normalized and enriched for
						full-length clones and was constructed by C. Gruber
						(Invitrogen). Research Genetics tracking code 019. Note:
						this is a NIH-MGC library."
BASE COUNT	207	a	220	g	184	t
ORIGIN						
Query Match	37.3%	Score 760.8;	DB 13;	Length 799;		
Best Local Similarity	99.4%;	Pred. No. 2.9e-142;				
Matches 795;	Conservative %	0;	Mismatches 2;	Indels 3;	Gaps 3;	
Oy	12	TGGGCGTCGGTGCGGCTCCAGACAGAGTTCTGTGGGAGTGGGCTCAGGGGCCCTGCAC	71			
Db	1	TGGCGTCGGTGCGGCTCCAGACAGAGTTCTGTGGGAGTGGGCTCAGGGGCCCTGCAC	60			
Oy	72	GGACCCCTACCTCTTGTAATGATTCTCGGGGGCCACGGGACCAGCAATCACAGCTGGCGT	131			
Db	61	GGACCCCTACCTCTTGTAATGATTCTCGGGGGCCACGGGACCAGCAATCACAGCTGGCGT	120			
Oy	132	TGCAGCTAAGGCGGCGGCTGGGGGTGGAGTGTGACGGCTGACTCCATGCAGGTCTATG	191			
Db	121	TGCAGCTAAGGCGGCGGCTGGGGGTGGAGTGTGACGGCTGACTCCATGCAGGTCTATG	180			
Oy	192	AAGGCTGAACATCATCAACAACAGCTTTCTGCCCAAGACAGACAGATTCGCCGCACC	251			
Db	181	AAGGCTGAACATCATCAACAACAGCTTTCTGCCCAAGACAGACAGATTCGCCGCACC	240			
Oy	252	ACATGATCAGCTTGTGGATCCTCTGTGACCAATTACAGAGTGATGCATTGAGAATA	311			
Db	241	ACATGATCAGCTTGTGGATCCTCTGTGACCAATTACAGAGTGATGCATTGAGAATA	300			
Oy	312	GAGCAACTGCTGTGATTAAGATATATTTGCCCGAGCAAAATTCATTTGTTGTGGAG	371			
Db	301	GAGCAACTGCTGTGATTAAGATATATTTGCCCGAGCAAAATTCCTA-TGTTGGGAG	359			
Oy	372	GACCAATTATTACATTNAATCTCTGCTCTGGAAAGTCTTGTCAAATCAACAGCCCACG	431			
Db	360	GACCAATTATTACATTNAATCTCTGCTCTGGAAAGTCTTGTCAAATCAACAGCCCACG	419			
Oy	432	AGATGGGCACTGAGAAAGTAGATTGACCGAAAAGTGGAGCTTGAAAAAGAGATGGCTTG	491			
Db	420	AGATGGGCACTGAGAAAGTAGATTGACCGAAAAGTGGAGCTTGAAAAAGAGATGGCTTG	479			
Oy	492	TACTTCACAAACGCCTAAGCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCATG	551			
Db	480	TACTTCAC -AACGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCATG	538			
Oy	552	ACAAACGCAAGGTGGCCAGGAGCTTGCAAGTTTTTGAAAGAAACAGAGATTCCTATATG	611			
Db	539	ACAAACGCAAGGTGGCCAGGAGCTTGCAAGTTTTTGAAAGAAACAGAGATTCCTATATG	598			
Oy	612	AATTCTCCTCATGCTAAACATACGAGAAGGTGGTGGCCCTTGAGAGTCTCTGAAAGT	671			
Db	599	AATTCTCCTCATGCTAAACATACGAGAAGGTGGTGGCCCTTGAGAGTCTCTGAAAGT	658			
Oy	672	TCTCTAACCCCTTGACCTTTGGCTTCATGCTGACACGAGCTCTAGATGAGCGCTTG	731			
Db	659	TCTCTAACCCCTTGACCTTTGGCTTCATGCTGACACGAGCTCTAGATGAGCGCTTG	718			
Oy	732	ATAAGAGGCTGGAGACATGCTGTGCTGGGCTT -GGAGGAATCAAGATTTTCAC	790			
Db	719	ATAAGAGGCTGGAGACATGCTGTGCTGGGCTTGGAGGAATCAAGATTTTCAC	778			
Oy	791	AGAGCTTAATACAGAAAGAA 810				
Db	779	AGAGCTTAATACAGAAAGAA 798				

OY	962	GTCATCGCTTGAAGCTATCAGTAGTCTGCAAGTGGGAGAGTGCTGTCTTGAACCTGCT	1021
Db	964	AACCTAGAGATATGCCCGGAAACAAGACCAGGAGTTTAAAAACCCCTTTTGGAG- ----	1018
OY	1022	CTTGAATCGTGCAAAAGTTTTCATCCAGGGGCCCAAGCCCTACAGCCACTCATATAAGATG	1081
Db	1019	----- -AAGTGCACAAACCTACAGCCATCGCAATGTAAGT	1052
OY	1082	CCATCATATGAAGCTGAGAACAAGAAAGTTATCACCTGTGTGACCTCTGTGATCGAATC	1141
Db	1053	GCATTCATATGAAGTGAACAAAGAAAGTTAACCATATGTGTGACCTCTGTGTACCCGGATC	1112
OY	1142	ATCATTTGGGGATCGGGAATGGGACACGCAATATAATCCAATCCCACCTGGAACACATG	1201
Db	1113	ATCATTTGGGGACCGGGAATGGGACACCAATTATGAATCCAATCTCACTTGCACCAATTG	1172
OY	1202	AAGAAAAGAGAAGATTTGGACTCAGATCTGTCAACACCATAGAAAGTCAGAGTTTCC	1261
Db	1173	AAGAAAAGAGAAGTTTGGACTTATGACCTGTCACTGTGCACAGAAAGTCAAAGTAAATCC	1232
OY	1282	CCAGCATATTAACAAGAACTTAAGGGAAGGATCCCCAGGGCACAATATGATTAACAGTGG	1321
Db	1233	CCAGCATGTGACCCGGAAACGCAATCGAGGGGGAATCTCGGGGACACACAATTAACAGCTG	1292
OY	1322	AAATCAGCAGTTTAAAGACATGTCAAGTGGCTTTGGAAAGGTGGTGGAGATCCAGTTTC	1381
Db	1293	AAGGCCATGTTTGAAGACATGCTATGAGCTTATGGCCTTTGCAGA - GAGTGGGGAATCAAGCC	1351
OY	1382	AGGAGGAGAGGGTATGTTTTCTGCCA - GTCTGGSCAAAAGAGTCTATCGGAAATCTTC	1440
Db	1352	AGGAGGGAGGGAGTGTGCTCTCCACAGCCCTGGAGCTGAGAAATCTTGGC - -----	1402
OY	1441	TGCATAGCAAGAAAGCTCCACATTTCTTTTGAATGTGGTTTAAAGTCTACGTTCTC	1500
Db	1403	----- -AGAAAGCCCCACCATCTTCTTTCATTTCTG- ----- -TGTGTGGTC	1441
OY	1501	TATAATAGAAACAGCAGGTCTTGTGACGCTCTTGTGTGGCTGATGTGTGTAATGATG	1560
Db	1442	TGCAGTGGAAACAGCAGGCTT - TCAGTCTCTGTGTGTGCTGTGTGTGTGTGTATGATG	1500
OY	1561	TAGTTCAGAAAGCAATTTTTTTTCTTTTGAACCTTAAAGCTTATATTTAAAGCAGC	1620
Db	1501	TAGTTCAGAGTGGG - - - TTTTTTCTTTGAACCTTAAAGGTTTATTTATTAACAGAGC	1557
OY	1621	ACAGATATCCACATTTTATATACATGAGGATCTTGTGGTGAATTAACAGAGATGTGCTGC	1680
Db	1558	ACAGATATCCACATTTTCTACTGAGGATCTTTTATGATGTGTAAATTAACAGATTAAGTGC	1617
OY	1681	ATCCCTTTAAAAAGATTTTATGTCCCTGACCTGTGCTAAATTTATATTAATTCAGATG	1740
Db	1618	AT - CTTTAAAAAGACGTTCTGTGCCCTGGGCTGTGCTAAATAATATGCTGTCCAGATG	1676
OY	1741	CTTTTGTGATGACTGGAAGTATTTGTGACCCACATATTTGGAGTCTTACATTTGAGTGA	1800
Db	1677	CTTTTGTGATGACTGGAAGTATTTGAAGCCGACGACAGAGCTGTGGACTCGAGAAACGC	1736
OY	1801	TGGCAGGAAGAGGCCATCTCCATTAAGATGATTAAGTGAACCAACTGTGTGGAATTT	1860
Db	1737	AGAAAGGAGTAC - ----- TGACAGAGATGATTAAGCAAACTCTCCAGCTGTGGAAT -	1789
OY	1861	CTACAGAGAAGAGGAATCAGACTGAGAGAAAGCTGTGACATAGGACTTGAAGAC - CAAG	1919
Db	1790	--- CTACAGAGAAGAGGGGTGACAGGCTGAGAGTGTGTGACCTGGAATCTTGAAGACATAAAG	1846
OY	1920	ACTTTGAATTTTGCAGCTGCTCATGTG	1947
Db	1847	ACACAGAGTCTGCTGAGCTGCTCYGTG	1874

DEFINITION	AL530574 LTI_NFL001.NBC4 Homo sapiens CDNA clone CS0DD008YN11 3 prime, mRNA sequence.
ACCESSION	AL530574
VERSION	AL530574.1 GI:12794067
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1, W.B., Gruber, C., Jesse, J. and Polayes, D.
TIITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: <a href="mailto:seque@genoscope.cns.fr">seque@genoscope.cns.fr</a> , Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> .
FEATURES	Location/Qualifiers
source	1..987
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="CS0DD008YN11"
	/clone_1lb="LTI_NFL001.NBC4"
	/sex="male"
	/tissue.type="neuroblastoma cells"
	/lab_host="DH10B"
	/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-G119(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9600 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : <a href="mailto:filiang@life.com">filiang@life.com</a> URL : <a href="http://fulllength.invitrogen.com">http://fulllength.invitrogen.com</a> "
BASE COUNT	281 a 230 c 180 g 284 t 12 others
ORIGIN	
Query Match	46.1%; Score 941.4; DB 9; Length 987;
Best Local Similarity	97.2%; Pred. No. 2e-178;
Matches 956; Conservative 10; Mismatches 17; Indels 1; Gaps 1.	
Oy	1019 GCTCTTGAATGTGTCGAAGTTTCATCCAGGGCCACACAGCTACAGCCACTCCATAAG 1078
Db	987 GCTCTTGAATGTGTCGAAGTTTCATCCAGGGCCACACAGCTACAGCCACTCCATAAG 928
Oy	1079 ATGCATCAATGAAGCTGAGAACGAAGAAGTATTCACCTGTGTGACCTGTGTATCGA 1138
Db	927 ATGCATCAATGAAGCTGAGAACGAAGAAGTATTCACCTGTGTGACCTGTGTATCGA 868
Oy	1139 ATCATCATTTGGGGATCGCGAATGGGCGAGCGACATTAATCCCAATCCCATTTGAACAA 1198
Db	867 ATCATCATTTGGGGATCGCGAATGGGCGAGCGACATTAATCCCAATCCCATTTGAACGA 808
Oy	1199 CTGAAGAAAGAGAGATTGGACTCAGTCTGTCTCAACCATGAAGAAAGTCAGAGTGT 1258
Db	807 CTGAAGAAAGAGAGATTGGACTCAGTCTGTCTCAACCATGAAGAAAGTCAGAGTGT 748
Oy	1259 TCCCGACACTATACAAAGAACTTAAGGAAGGATCCCGAGGCGACGAATATCAAG 1318
Db	747 TCCCGACACTATACAAAGAACTTAAGGAAGGATCCCGAGGCGACGAATATCAAGG 688
Oy	1319 CTGAATCAGAGCTTTAAGAGACATGTCACAGTGGCTTTGGAAGGTGGTGGGATCCAG 1378
Db	687 CTGAATCAGAGCTTTAAGAGACATGTCACAGTGGCTTTGGAAGGTGGTGGGATCCAG 628
Oy	1379 TTCAGAGGAGAGGGGTATGTTGTCTCCAGTCTGGCCAAAGAGTGCATCGGAATTC 1438
Db	627 TTCAGAGGAGAGGGGTATGTTGTCTCCAGTCTGGCCAAAGAGTGCATCGGAATTC 568
Oy	1439 TCTGCATGACGAGAAAGCTCCCACTTTCTTTGATGTGTTTTAAAGTTCACGTTTC 1498



QY 548 CATGCAAAACGCAAGTGGCCAGAGCTTCAGATTTTGAAGAAACAGAACTCTCTCAT 607  
 |||||  
 Db 552 CATGCAAAACGCAAGTGGCCAGAGCTTCAGATTTTGAAGAAACAGAACTCTCTCAC 611  
 QY 608 AGTGAATTTCTCCATCTCAACATACGAGAAAGAGTGTGTGCTCCCTTGAGAGTCTG 667  
 |||||  
 Db 612 AGTGAATTTCTCCATCTCAACATACGAGAAAGAGTGTGTGCTCCCTTGAGAGTCTG 671  
 QY 668 AAGTTCTCTAACCTTCCTGATCTTGGCTTCATGCTGACCAAGGAGTTCTAGATGAGCC 727  
 |||||  
 Db 672 AAGTTCTCTAACCTTCCTGATCTTGGCTTCATGCTGACCAAGGAGTTCTAGATGAGCC 731  
 QY 728 TTGATTAAGAGGTTGATGATCTGCTGCTGCTGCTTGGAGAACTAAGATTTT 787  
 |||||  
 Db 732 TTGATTAAGAGGTTGATGATCTGCTGCTGCTGCTTGGAGAACTAAGATTTT 791  
 QY 788 CACAGACCTATATATCAGAAATGTTTCGAAATATAGCCAGAGCTATCAATGATGATC 847  
 |||||  
 Db 792 CACAGACCTATATATCAGAAATGTTTCGAAATATAGCCAGAGCTATCAATGATGATC 851  
 QY 848 TTCCATCAATTTGCTTCAAGGAATTTCAAGATCTGATCACTGAGGGAATATGACACA 907  
 |||||  
 Db 852 TTCCATCAATTTGCTTCAAGGAATTTCAAGATCTGATCACTGAGGGAATATGACACA 911  
 QY 908 CTGGAGACTAGTAACCACTTTTAAAGAAAG----- 938  
 |||||  
 Db 912 CAGAGACTAGTAACCACTTTTAAAGAAAGTATGAGGCTCTGAAACAACTAATAAG 971  
 QY 938 -----GACCTGCTCC 949  
 Db 972 AGATATGCCCCGAAACAGACCCGATGGTTTAAACCCGCTTTTGGACAGACTGGGCC 1031  
 QY 950 ATTGTCCTCCCTGTATATGCTTGAAGATATCTGATGCTTCAGATGGAGAGCTGTT 1009  
 |||||  
 Db 1032 AGTGTCCCTCCGATATATGCTTGAAGATATCTGATGCTTCAGATGGAGAGCTGTT 1091  
 QY 1010 CTGGAACCTGCTTGAATATCTGCAAGTTTCAATCCAGGCGCACAGGCTTACAGCCACT 1069  
 |||||  
 Db 1092 CTGGAACCTGCTTGAATATCTGCAAGTTTCAATCCAGGCGCACAGGCTTACAGCCACT 1151  
 QY 1070 CCAATTAAGATGCTATACATGATGATGAGCAAGAGATTAACACTGCTGAGCC 1129  
 |||||  
 Db 1152 CCAATTAAGATGCTATACATGATGAGCAAGAGATTAACACTGCTGAGCC 1211  
 QY 1130 TGTGATGCAATCATATTTGGGATCGCAATGCGGACACATTAATCCAAATCCAC 1189  
 |||||  
 Db 1212 TGTGATGCAATCATATTTGGGATCGCAATGCGGACACATTAATCCAAATCCAC 1271  
 QY 1190 TTGAACCAACTGAAAGAAAGAAAGATTTGACATGCTGCTCAACACCTATAGAACT 1249  
 |||||  
 Db 1272 TTGAACCAACTGAAAGAAAGAAAGATTTGACATGCTGCTCAACACCTATAGAACT 1331  
 QY 1250 CAGAGTGTTCCTCCAGACTATATCAAAAGAACTTAAGGGAAGGATCCGAGGAGAT 1309  
 |||||  
 Db 1332 CAGAGTGTTCCTCCAGACTATATCAAAAGAACTTAAGGGAAGGATCCGAGGAGAT 1391  
 QY 1310 GATCAAGAGCTGAAATGAGGCTTTAAGAGATGCTCCAGTGGCTTTGAAAGGTG 1369  
 |||||  
 Db 1392 AATCAAGAGCTGAAATGAGGCTTTAAGAGATGCTCCAGTGGCTTTGAAAGGTG 1450  
 QY 1370 GGAATCCAGTTTCAAGAGGAGGAGGATTTTGTCTCCCA-GTCTGGCAAGAGAGTCTA 1428  
 |||||  
 Db 1451 GGAATCCAGTTTCAAGAGGAGGAGGATTTTGTCTCCCA-GTCTGGCAAGAGAGTCTA 1510  
 QY 1429 TGCGGAATTTCTGATAGCAAGAAAGTCCACCACTTTTGTGATGCTGTTTAAAG 1488  
 |||||  
 Db 1511 GGC-----AGAGGCCCCACCATCTTCTTTATTTCTGTGCT----- 1546  
 QY 1489 TCTCAGTTCTCTATATAGAAACAGAGGCTTGTGAGCTCTTGTGCTGATGCTGT 1548  
 |||||  
 Db 1547 -----GTGCTCTGCTGAGTGAAGAGAGGCTT-TCAAGTCTCTGTGCTGTGCTGT 1599  
 QY 1549 CTGGAATGATGATGCTGAGGAAGCAATTTTTTTTCTTTGACCTTAAAGGTTCTAT 1608

Db 1600 CTGATATATGATGATGCTAG-----AGTGGATTTTTTTTCTTGAACCTTAAAGGTTTAT 1656  
 QY 1609 ATTAAGAGCAGACAGATTTCCATATTTATACATGAGATCTTCTTGGGATATAC 1668  
 |||||  
 Db 1657 TTGAAGAGGAGCAGATGCGACATTTTCTACTGAGATCTTTTATGAGGATATAC 1716  
 QY 1669 AGATTTGACTGATCCCTTTTAAAGAGATTTTATGCTCCCTGACTGCGCTTAAATATCT 1728  
 |||||  
 Db 1717 AAGATTCAGTGCAT-CTTTTAAAGAGCGTCTTGTCCCTGGGCTGGCTTAAATATAGCT 1775  
 QY 1729 AATTTCAGATCTTTTGTAGATGATGATGATGATGATGATGATGATGATGATGAT 1788  
 |||||  
 Db 1776 CGTTCCAGATCTTTTGTAGATGATGATGATGATGATGATGATGATGATGATGAT 1835  
 QY 1789 GATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848  
 |||||  
 Db 1836 CTGAGAGAGGAGCAGAAAGAGATG-----TGCAGAGAGATGATTAAGCAAACTCTCCA 1889  
 QY 1849 GTTCTCGAATTTCTACAGAGAGAGGAGATCAGACTGAGAGAGCTGATGATGATGAT 1908  
 |||||  
 Db 1890 GCTCTGTGAT-----CTACAGAGAGAGGAGGAGTGAAGTGTGATGATGATGATGAT 1945  
 QY 1909 GAAGAC-CAAGACTTTGAAATTTGCGAGCTGCTCATGTC 1947  
 |||||  
 Db 1946 GAAGACTAAAGAGAGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1985  
 RESULT 2  
 AK010176  
 LOCUS  
 DEFINITION  
 Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310075G14:homolog to TRNA ISOPTERYLYPHROPHOSPHATE TRANSFERASE, full insert sequence.  
 AK010176  
 VERSION  
 AK010176.1 GI:12845432  
 KEYWORDS  
 HTC; CAP trapper.  
 SOURCE  
 Mus musculus (strain: C57BL/6J) adult male tongue cDNA to mRNA, clone:1b:RIKEN full-length enriched mouse cDNA library  
 clone:2310075G14.  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE  
 1 Carninci, P., and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE  
 99279253  
 PUBMED  
 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE  
 20499374  
 PUBMED  
 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunaka, T., Yoshino, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsui, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multiplexed sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE  
 20530913  
 PUBMED  
 11076861  
 4  
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:27:18 ; Search time 3842.56 Seconds  
(without alignments)  
8602.339 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041  
Sequence: 1 CTGCCATAGATGCGTCGCG.....TTTACAGAGAAAAA 2041

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthm:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1128.4	55.3	2061	11	AK003556 Mus muscu
2	1069.6	52.4	1951	11	AK010176 Mus muscu
3	941.4	46.1	987	9	AL530574 AL530574
4	760.8	37.3	799	13	AL530574 AL530574
5	733.6	35.9	1088	14	BM800217 BM800217
6	727.4	35.6	743	14	BM975520 UI-CF-EN1

Result No.	Score	Query Match	Length	ID	Description
7	690.6	33.8	725	10	BE315223 BE315223
8	688.2	33.7	725	9	AI133396 AI133396
9	677.8	33.2	724	14	BM721352 UI-E-EO1-
10	666.2	32.6	869	12	BM721352 UI-E-EO1-
11	664.8	32.6	957	13	BM721352 UI-E-EO1-
12	662.8	30.8	772	12	BM721352 UI-E-EO1-
13	606.8	29.7	624	14	BM721352 UI-E-EO1-
14	599.8	29.4	711	12	BM721352 UI-E-EO1-
15	594.8	29.1	783	13	BM721352 UI-E-EO1-
16	594.8	29.1	783	13	BM721352 UI-E-EO1-
17	589	28.9	601	9	AI207688 AI207688
18	578.2	28.3	922	9	AI207688 AI207688
19	574.4	28.1	578	14	BM847727 K-EST0127
20	570	27.9	643	10	BM847727 K-EST0127
21	569.6	27.9	758	13	BM847727 K-EST0127
22	566.8	26.8	747	12	BM847727 K-EST0127
23	546.4	26.8	559	9	AI178450 AI178450
24	544.2	26.7	716	13	AI178450 AI178450
25	541.8	26.5	806	12	BE965936 BE965936
26	539.4	26.4	724	12	BE965936 BE965936
27	535	26.2	554	9	AI1749267 AI1749267
28	530.4	26.0	543	9	AI1749267 AI1749267
29	526.2	25.8	836	12	BM757745 BM757745
30	524.4	25.7	562	12	BM757745 BM757745
31	515.4	25.3	528	12	BF064275 BF064275
32	511.2	25.0	963	13	BI462398 BI462398
33	509.2	24.9	645	12	BC288423 BC288423
34	504.8	24.7	946	10	BB656365 BB656365
35	500.4	24.5	513	10	AM104415 AM104415
36	495.8	24.3	510	10	AM263011 AM263011
37	490.2	24.0	518	9	AI459111 AI459111
38	487.4	23.5	500	9	AI192804 AI192804
39	476	23.3	490	9	AI127851 AI127851
40	470.4	23.0	487	9	AI337829 AI337829
41	466.8	22.9	491	12	BF433525 BF433525
42	465.4	22.8	546	10	AV607062 AV607062
43	459	22.5	593	10	AM444470 UI-H-B13-
44	458.6	22.5	862	12	BF142859 BF142859
45	458.6	22.5	676	10	BE614088 BE614088

## ALIGNMENTS

RESULT 1  
AK003556  
LOCUS  
DEFINITION  
Mus musculus 18 days embryo whole body cDNA, RIKEN full-length  
enriched library, clone:110007017, homolog to TRNA  
ISOPENTENYL-PIROPHOSPHATE TRANSFERASE, full insert sequence.  
ACCESSION  
AK003556  
VERSION  
AK003556.1 GI:12834290  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (strain: C57BL/6J) 18 days embryo cDNA to mRNA,  
clone:110007017.  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
PUBMED  
99279253  
10349636  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subcloning of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
JOURNAL  
MEDLINE  
PUBMED  
20499374  
11042159

NUMBER OF SEQUENCES: 1  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Amster, Rothstein & Ebenstein  
 STREET: 90 Park Avenue  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10016  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Word Processor (ASCII)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/075,904  
 FILING DATE: May 11, 1998  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/247,901  
 FILING DATE: May 23, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bogosian, Elizabeth A  
 REGISTRATION NUMBER: 39,911  
 REFERENCE/DOCKET NUMBER: 96700/475  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 697-5995  
 TELEFAX: (212) 286-0854 or 286-0082  
 TELEX: TWX 710-581-4766  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 50341  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE:  
 DESCRIPTION: L5 shuttle phasmid sequence  
 HYPOTHEICAL: NO  
 ANTI-SENSE:  
 FRAGMENT TYPE:  
 ORIGINAL SOURCE:  
 ORGANISM: L5 mycobacteriophage  
 STRAIN:  
 INDIVIDUAL ISOLATE:  
 DEVELOPMENTAL STAGE:  
 HAPLOTYPE:  
 TISSUE TYPE:  
 CELL TYPE:  
 CELL LINE:  
 ORGANELLE:  
 IMMEDIATE SOURCE:  
 POSITION IN GENOME:  
 CHROMOSOME/SEGMENT:  
 FEATURE:  
 NAME/KEY:  
 LOCATION:  
 IDENTIFICATION METHOD:  
 OTHER INFORMATION:  
 PUBLICATION INFORMATION: No. 5994137e  
 AUTHORS:  
 TITLE:  
 JOURNAL:  
 VOLUME:  
 PAGES:  
 DATE:  
 DOCUMENT NUMBER:  
 FILING DATE:  
 PUBLICATION DATE:  
 RELEVANT RESIDUES IN SEQ ID NO:

```

Query Match      1.8%;   Score 36.4;   DB 2;   Length 50341;
Best Local Similarity 49.0%;   Pred. No. 10;
Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

```

QY	40	TCGTGGGAGTGGGCTCAGGGGGCCCTGCACAGGAGCCCTACCTCTTTGATGAGTATCTGG	99
Db	2478	TCCTCTGTGATGTTCTCTGTGCGCGGGCTTGAGCCGGCCCACTACTCTTTGCCGACCTGCG	24119
QY	100	GGCCACAGGGGACCGGCAATATCAACGCTGGCGTTGCAGCTAGAGCCACGGGCTGGCGGTGA	159
Db	2418	GGCGAATTCAGGCTCTGACCTCTGCTGGCCCGCTTGCGGCGTTCGGCGAGCTTCCTGTGGAGA	23599
QY	160	GATCGTGTAGCGCTGACTCCATGTCAGGTCTATATAGAGCCCTAGACATCATCACCACAAAGT	219
Db	2358	CTCTGAGGCTCCCTGGACCCCTCGTATGTAGATGATCTCAGAGACCTTCGGCACGCGCCCACT	22999
QY	220	TTTCGCCCAAGACAGAG	237
Db	2298	TCGTGGCCGACAGCCGGG	2281

Search completed: April 21, 2003, 16:30:04  
Job time : 3773.69 secs



STATE: CA  
COUNTRY: US  
ZIP: 91320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/053,374A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: COOK, ROBERT R.  
REGISTRATION NUMBER: 31,602  
REFERENCE/DOCKET NUMBER: A-514  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1689 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 132..971  
US-09-053-374A-4

Query Match 1.8%; Score 36.4; DB 4; Length 1689;  
Best Local Similarity 66.7%; Pred. No. 1.7;  
Matches 52; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1964 CTGCTTTCTTGTAGTACAACTATTTTATTGTAAGTTTAAATGAAGAAAAATT 2023  
DB 1544 CTACATTTCTATGTGTGAGGCTTTATTATTATTTTAAATTAATTAATATTATTA 1603  
QY 2024 TACAAGAAAAAATAA 2041  
DB 1604 AAATGAAAAAATAA 1621

RESULT 14  
US-08-247-901C-1/C  
Sequence 1, Application US/08247901C  
Patent No. 5750384  
GENERAL INFORMATION:  
APPLICANT: Jacobs et al  
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amster, Rothstein & Edenstein  
STREET: 90 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: word processor (ASCII)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,901C  
FILING DATE: May 23, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/057,531  
FILING DATE: April 29, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Bogosian, Elizabeth A  
REGISTRATION NUMBER: 39,911  
REFERENCE/DOCKET NUMBER: 96700/273  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 697-5995  
TELEFAX: (212) 286-0854 or 286-0082  
TELEX: TWX 710-581-4766  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50341  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: L5 shuttle phasmid sequence  
HYPOTHETICAL: No  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: L5 mycobacteriophage  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE:  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION: No. 5750384e  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-08-247-901C-1

Query Match 1.8%; Score 36.4; DB 1; Length 50341;  
Best Local Similarity 49.0%; Pred. No. 10;  
Matches 97; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 40 TCCTGTGGGAGTGGGCTCAGAGGGGCTGCAAGGACCTTCTTGTAGTATTCGG 99  
DB 2478 TCTCTGTGTAGTCTCTGTGCGCGGCTGAGCGGCCAGTCACCTTGTGCGACTCGC 2419  
QY 100 GGGCAGGGGCAACGGCAATCCAGCGCTGCGAGTACAGCCGCGGGGCGGA 159  
DB 2418 GGGCAAGTTCAGCCTCGACCTGCTGCGCCCTGGGCTTCGCGAGCTTCTGTGGAGA 2359  
QY 160 GATGTCAGCGCTGCTCATGAGGCTATGAGGCGCTAGACATCATCACCACAAAGT 219  
DB 2358 CCTGAGCGTCTCCGAGACCCGCTGAGAGCTGTGAGAGATCTGAGAGACCCGCGACCGCCCACT 2259  
QY 220 TTCTGCCCAAGACAGAG 237  
DB 2298 TCGTGGCCGACAGCGGG 2281

RESULT 15  
US-08-075-904-1/C  
Sequence 1, Application US/09075904  
Patent No. 5994137  
GENERAL INFORMATION:  
APPLICANT: Jacobs, et al.  
TITLE OF INVENTION: L5 SHUTTLE PHASMIDS

```

; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

Query Match 1.8%; Score 37.4; DB 4; Length 2346;  
 Best Local Similarity 64.4%; Pred. No. 1;  
 Matches 56; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

```

QY 1955 TTTATCTGCTGCTTTCTATAGTTACAATCTATATTTTATGACGTTAAATAA 2014
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2197 TTTTATGATGATATTTGTTATTTCCAGCAATTAATTTGTAAGTTGCAAAAAA 2256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2015 GAAAAAATTACAGAAAAA 2041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2257 AAAAAAAAAAAAAAAAAA 2283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12  
 US-07-847-010-12  
 ; Sequence 12, Application US/07847010  
 ; Patent No. 5693495  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Breiteneder, Helmo  
 ; APPLICANT: Reikstorfer, Arnold  
 ; APPLICANT: Valenta, Rudolf  
 ; APPLICANT: Hoffmann, Sommergruber, Karin  
 ; APPLICANT: Breitenbach, Michael  
 ; APPLICANT: Kraft, Dietrich  
 ; APPLICANT: Rumpold, Helmut  
 ; APPLICANT: Scheiner, Otto

```

; APPLICANT: Ebner, Christof
; APPLICANT: Ferreltra, Fatima
; TITLE OF INVENTION: Allergens of Alder Pollen and
; TITLE OF INVENTION: Applications Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/847,010
; FILING DATE: 01-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jones III, Harry C
; REGISTRATION NUMBER: 20,280
; REFERENCE/DOCKET NUMBER: 6530-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 742 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: hazel (Corylus sp.)
; IMMEDIATE SOURCE:
; LIBRARY: POLLEN FROM ALLERGEN AB, ENGELHOLM, SWEDEN
; US-07-847-010-12

```

Query Match 1.8%; Score 36.6; DB 1; Length 742;  
 Best Local Similarity 56.1%; Pred. No. 0.94;  
 Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

```

QY 1919 GACCTTGAATTTGCGAGCTGCTCATGTGAGTTATATACACTGCTCTTCTATTGA 1978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 604 GATGTTTAACTCGGAAGAAATGCAAACTGAATGTTGTAATTTTGCATATATACAA 663
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1979 GTTACAATCTAATATTTTATTTGAGTTAAATTAAGAAAAAATTTACAGAAAAA 2038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 664 TAATGGAAGATPAACATCATTTGAAGTTCAAAAAAAAAAAAAAAAAAAAAAAAAA 723
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2039 AAA 2041
    |||
DB 724 AAA 726
    |||

```

RESULT 13  
 US-09-053-374A-4  
 ; Sequence 4, Application US/09053374A  
 ; Patent No. 6462177  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YEN, KWANG-MU  
 ; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: AMGEN INC.  
 ; STREET: ONE AMGEN CENTER DRIVE  
 ; CITY: THOUSAND OAKS

EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043, 6658
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043, 312
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043, 3133
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043, 6727
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/043, 315
EARLIER FILING DATE:	1997-04-11
EARLIER APPLICATION NUMBER:	60/048, 9744
EARLIER FILING DATE:	1997-06-06
EARLIER APPLICATION NUMBER:	60/056, 8866
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 8777
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 8899
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 8933
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 630
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 8828
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 6373
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 9030
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 8888
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 8797
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 880
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 845
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 874
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 910
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 864
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 6313
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 845
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/056, 892
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/057, 761
EARLIER FILING DATE:	1997-08-22
EARLIER APPLICATION NUMBER:	60/047, 599
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047, 588
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047, 585
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047, 586
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047, 590
EARLIER FILING DATE:	1997-05-23
EARLIER APPLICATION NUMBER:	60/047, 590
EARLIER FILING DATE:	1997-05-23

EARLIER APPLICATION NUMBER: 60/043, 6711
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 6659
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043, 315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048, 974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056, 886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056, 892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057, 761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047, 595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047, 586

```

1 EARLIER FILING DATE: 1997-05-23
2 EARLIER APPLICATION NUMBER: 60/047,590
3 EARLIER FILING DATE: 1997-05-23
4 EARLIER APPLICATION NUMBER: 60/047,594
5 EARLIER FILING DATE: 1997-05-23
6 EARLIER APPLICATION NUMBER: 60/047,589
7 EARLIER FILING DATE: 1997-05-23
8 EARLIER APPLICATION NUMBER: 60/047,593
9 EARLIER FILING DATE: 1997-05-23
10 EARLIER APPLICATION NUMBER: 60/047,614
11 EARLIER FILING DATE: 1997-05-23
12 EARLIER APPLICATION NUMBER: 60/043,578
13 EARLIER FILING DATE: 1997-04-11
14 EARLIER APPLICATION NUMBER: 60/043,576
15 EARLIER FILING DATE: 1997-04-11
16 EARLIER APPLICATION NUMBER: 60/047,501
17 EARLIER FILING DATE: 1997-05-23
18 EARLIER APPLICATION NUMBER: 60/043,670
19 EARLIER FILING DATE: 1997-04-11
20 EARLIER APPLICATION NUMBER: 60/056,632
21 EARLIER FILING DATE: 1997-08-22
22 EARLIER APPLICATION NUMBER: 60/056,664
23 EARLIER FILING DATE: 1997-08-22
24 EARLIER APPLICATION NUMBER: 60/056,876
25 EARLIER FILING DATE: 1997-08-22
26 EARLIER APPLICATION NUMBER: 60/056,881
27 EARLIER FILING DATE: 1997-08-22
28 EARLIER APPLICATION NUMBER: 60/056,909
29 EARLIER FILING DATE: 1997-08-22
30 EARLIER APPLICATION NUMBER: 60/056,875
31 EARLIER FILING DATE: 1997-08-22
32 EARLIER APPLICATION NUMBER: 60/056,862
33 EARLIER FILING DATE: 1997-08-22
34 EARLIER APPLICATION NUMBER: 60/056,887
35 EARLIER FILING DATE: 1997-08-22
36 EARLIER APPLICATION NUMBER: 60/056,908
37 EARLIER FILING DATE: 1997-08-22
38 EARLIER APPLICATION NUMBER: 60/048,964
39 EARLIER FILING DATE: 1997-06-06
40 EARLIER APPLICATION NUMBER: 60/057,650
41 EARLIER FILING DATE: 1997-09-05
42 EARLIER APPLICATION NUMBER: 60/056,884
43 EARLIER FILING DATE: 1997-08-22
44 EARLIER APPLICATION NUMBER: 60/057,669
45 EARLIER FILING DATE: 1997-09-05
46 EARLIER APPLICATION NUMBER: 60/049,610
47 EARLIER FILING DATE: 1997-06-13
48 EARLIER APPLICATION NUMBER: 60/061,060
49 EARLIER FILING DATE: 1997-10-02
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
```

```
Patent No. 6465229
GENERAL INFORMATION:
APPLICANT: Rafalski, Antoni J.
APPLICANT: Fader, Gary M.
APPLICANT: Cahoon, Rebecca E.
TITLE OF INVENTION: Plant Caffeyol-CoA O-Methyltransferase
FILE REFERENCE: B01284 US NA
CURRENT APPLICATION NUMBER: US/09/452,239
CURRENT FILING DATE: 1999-12-01
EARLIER APPLICATION NUMBER: 60/110,594
EARLIER FILING DATE: 1998-December-02
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Microsoft Office 97
SEQ ID NO 21
LENGTH: 980
TYPE: DNA
ORGANISM: Glycine max
US-09-452-239-21

Query Match
Best Local Similarity 1.9%; Score 37.8; DB 4; Length 980;
Matches 66; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1929 TTTCGAGTCTCATGTCAGTATTAATCACTGCTCTTCTATTGATTACAATC 1988
DB 867 TTTCAGAGCTTGTGTGTTCTGTGTCAGCTAGAAATTTAGCTGAAGATCAATC 926

QY 1989 TAATATTTTATGAAGTTAAATTAAGAAAAATTTACAAGAAAAAAA 2041
DB 927 AACATTATCTTATTTAAAAAAAATTTAAAAAAAATTTAAAAAAA 979

RESULT 9
US-09-453-323-1
Sequence 1, Application US/09453323
GENERAL INFORMATION:
APPLICANT: Ye, Zheng-Hua
TITLE OF INVENTION: MODIFICATION OF LIGNIN CONTENT AND COMPOSITION IN
PLANTS
FILE REFERENCE: 235,00090101
CURRENT APPLICATION NUMBER: US/09/453,323
CURRENT FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: 60/110,676
PRIOR FILING DATE: 1998-12-02
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patentlin Ver. 2.0
SEQ ID NO 1
LENGTH: 1507
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-453-323-1

Query Match
Best Local Similarity 1.9%; Score 37.8; DB 4; Length 1507;
Matches 60; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1945 GTGTGAGTATTTATCACTGCTCTTCTATTCAGTATTAATTTATTTATGAG 2004
DB 1366 GTTGTAGCTGTATTCGCTGTTATTCCTCAATTCCTCCCTAAGCAAGATTTAGCAGATG 1425

QY 2005 TTTAAATTAAGAAAAATTTACAAGAAAAAAA 2041
DB 1426 ATAAAAAAAATTTAAAAAAAATTTAAAAAAA 1462

RESULT 10
US-09-149-476-316
Sequence 316, Application US/09149476
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
```

```
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
EARLIER FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,560
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
```



DB 4122 GGTGGAAACGAGATCATCGCGACTGCGAGAGATATACCGTGACATGACATCG 4181  
QY 207 TCACCAACAAGCTTCTGCGCCAGACGAG 235  
DB 4182 GCACAGCAAGACTTGCGCGACTATCAG 4210

## RESULT 6

US-08-974-691-7  
Sequence 7, Application US/08974691  
Patent No. 6225103

## GENERAL INFORMATION:

APPLICANT: Keisch, Gerald  
APPLICANT: Lin, Xinh  
TITLE OF INVENTION: Cloning and Characterization of Napsin  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center, 1201 W. Peachtree  
St.  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,691

FILING DATE: 20-NOV-1997

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/031,196

FILING DATE: 20-NOV-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/046,126

FILING DATE: 09-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: OMRF 166

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-873-8794

TELEFAX: 404-873-8795

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1910 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

US-08-974-691-7

Query Match 2.0%; Score 40.8; DB 4; Length 1910;  
Best Local Similarity 61.1%; Pred. No. 0.096;

Matches 66; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1934 GAGCTGCTATGTGTGATATATACATGCTCTCTTCTATGAGTTACAATCTATAT 1993  
DB 1796 GTGAACCGCATCGCGAGCGGTGTCGCGAGCTCCGTACCCAGTAATAATCCACTAT 1855  
QY 1994 TTTTATGAAGTTAATAAGAAAAATTACAGAAAAA 2041  
DB 1856 TTCCATTGAAAAA 1903

## RESULT 7

US-09-149-476-187  
Sequence 187, Application US/09149476  
Patent No. 6420526

## GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002pl

CURRENT APPLICATION NUMBER: US/09/149,476

FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,563

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,552

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,587

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,492

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,582

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,612

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,632

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043,580

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,568





QY 442 TGAGAAAGTATGACCGAATAAGAGGATGCTTGTACTTCA 498  
 DB 381 AATGAAACATTAATTAATTAAGATTGACACTTAATTAATTAAGCTTCA 440  
 QY 499 CAACGGCTAAGCAGGCTGACCCAGAAATGCTGACAGTGCATCCATGACAAAGC 558  
 DB 441 CGAATATTAGCTTCATCGACAAAGATGACCAAGATATACCTTAATACAGAAA 500  
 QY 559 CAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAA 599  
 DB 501 AAGAGTGTGCGACCAATAGATATTAATTTGAAAAACAAAA 541

## RESULT 2

US-09-103-840A-1/c  
 ; Sequence 1, Application US/09103840A  
 ; Patent No. 6294328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEISCHMAN, Robert D.  
 ; APPLICANT: WHITE, Owen R.  
 ; APPLICANT: FRASER, Claire M.  
 ; APPLICANT: VENTER, John C.  
 ; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
 ; FILE REFERENCE: 24366-20007.00  
 ; CURRENT APPLICATION NUMBER: US/09/103,840A  
 ; CURRENT FILING DATE: 1998-06-24  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 4411529  
 ; TYPE: DNA  
 ; ORGANISM: Mycobacterium tuberculosis  
 ; OTHER INFORMATION: H37Rv  
 US-09-103-840A-1

Query Match 2.9%; Score 58.4; DB 4; Length 4411529;  
 Best Local Similarity 53.7%; Pred. No. 5.1e-05;  
 Matches 175; Conservative 0; Mismatches 136; Indels 15; Gaps 2;

QY 87 TAGGATCTCTCGGGCCAGCGCAACCAATCCAGCTGCTGACGTAAGCCAGC 146  
 DB 3040756 TGCGGATTAATCGGACCGACCGCGCCGCAAGTCCCAATGCGCTGACCTGCCCGCCG 3040697  
 QY 147 GCGTCGGCGGT-----GAGATCGTCAAGCGCTGACTCCATGCAAGGTCTATGAAG 194  
 DB 3040696 GCGTCGGCGCCAGGCTGAGCGTGAATCGAAGCGCCGACGATGCAAGTCTATGCGC 3040637  
 QY 195 GCGTACATATCATCAACCAAGGTTTCTGCCCAAGAGCAGAGATCTGCCGGCACCA 254  
 DB 3040636 GCATGACATCGGACGCGCAAGTGCCTCAAGCGCGCGTGGCATCCCGCATCATC 3040577  
 QY 255 TGATGAGTTGTGATGCTCTGTGACCAATTAACAGTGGTGAATTCAGAAATAGAG 314  
 DB 3040576 AACTCGATGCTCTGATGATGACCAACAC---GGACCGTGGCGCTACCAAGCGGCGC 3040520  
 QY 315 CAACGTCTGTGATGAAGATATATTTGCCCGAGACAAATTCCTATTGTGTGGAGAA 374  
 DB 3040519 CCGGCGGAGCATGAGAGCGATCGACGCGCGGAGCGGTCCGCTGCTGTGGCGGCT 3040460  
 QY 375 CCAATTATTACATTGAATCTGCTC 400  
 DB 3040459 CGATGCTGTATGTCCAAATCCCTGCTC 3040434

## RESULT 3

US-08-232-463-14/c  
 ; Sequence 14, Application US/08232463  
 ; Patent No. 5670367  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DORNER, F.  
 ; APPLICANT: SCHEFFLINGER, F.  
 ; APPLICANT: FALKNER, F. G.

; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 1800 Diagonal Road, Suite 500  
 ; CITY: Alexandria  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22313-0299  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232,463  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/935,313  
 ; FILING DATE:  
 ; APPLICATION NUMBER: EP 91 114 300.6  
 ; FILING DATE: 26-AUG-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703)836-9300  
 ; TELEFAX: (703)683-4109  
 ; TELEX: 899149  
 ; INFORMATION FOR SEQ ID NO: 14:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 7218 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; CLONE: pTZptc-Fls  
 US-08-232-463-14

Query Match 2.6%; Score 52.2; DB 1; Length 7218;  
 Best Local Similarity 3.1%; Pred. No. 0.0001;  
 Matches 12; Conservative 222; Mismatches 155; Indels 0; Gaps 0;

QY 1073 ATAAAGATGCCATACATGAGCTGAGAAAGAGAGATATACCTGTGACCTCTGT 1132  
 DB 1449 AGAAGATTTGTGACRR 1390  
 QY 1133 GATCAATCATCATTTGGGATCGCAATGGCGACGCGCAATAAATCCAAATCCACTG 1192  
 DB 1389 RRR 1330  
 QY 1193 AACCAACTGAAGAAAGAAAGATTTGACTGATGCTGTCACACCATAGAAAGTCAG 1252  
 DB 1329 RRR 1270  
 QY 1253 AGTGTTCGCCAGCTATTAACAAAGACCTAAAGGAGGATCCCGAGCGCAATGAT 1312  
 DB 1269 RRR 1210  
 QY 1313 CAAGAGCTGAATGCAAGCTTTAAAGAGACATGCTCCAGTGGCGCTTGGAAAGTGGTGGG 1372  
 DB 1209 RRR 1150  
 QY 1373 ATCCAGTTGAGGAGGAGGATGTTGTCTCCAGTGGGCAAGAGAGTATGCG 1432  
 DB 1149 RRR 1090  
 QY 1433 GAATTCCTGATAGCAAGAAAGCTCCCA 1461  
 DB 1089 RRRRRRRRRRRRRRRRRRRRRRRATCGCA 1061

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:23:43 ; Search time 97.6922 Seconds

6407.139 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 2041

sequence: 1 CTGCCATTAAGATGGCGTCCG.....TTTACAAGAAAAAAA 2041

Scoring table: IDENTITY\_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum match 08

Listing first 45 summaries

Database :

```

1: /cgn2_6/ptodata1/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata1/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata1/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata1/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata1/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata1/1/ina/backfiles1.seq:*

```

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	65	3.2	582	4	US-09-134-001C-1334	Sequence 1334, App	
C	2	58.4	2.9	4411529	4	US-09-103-840A-1	Sequence 1, Appl	
C	3	52.2	2.6	7218	1	US-08-932-465-14	Sequence 14, Appl	
C	4	47.8	2.3	846	4	US-08-936-168A-65	Sequence 65, Appl	
C	5	46.6	2.3	5884	4	US-09-221-011B-497	Sequence 497, Appl	
C	6	40.8	2.0	1910	4	US-08-974-691-7	Sequence 7, Appl	
C	7	38	1.9	1605	4	US-09-149-476-187	Sequence 187, App	
C	8	37.8	1.9	980	4	US-09-453-239-21	Sequence 21, Appl	
C	9	37.8	1.9	1507	4	US-09-453-322-1	Sequence 1, Appl	
C	10	37.4	1.8	2025	4	US-09-149-476-316	Sequence 316, Appl	
C	11	37.4	1.8	2346	4	US-09-149-476-193	Sequence 193, App	
C	12	36.6	1.8	742	1	US-07-847-010-12	Sequence 12, Appl	
C	13	36.4	1.8	1689	4	US-09-053-374A-4	Sequence 4, Appl	
C	14	36.4	1.8	50341	1	US-08-247-901C-1	Sequence 1, Appl	
C	15	36.4	1.8	50341	2	US-09-075-904-1	Sequence 1, Appl	
C	16	36.4	1.8	52297	4	US-09-426-436-1	Sequence 1, Appl	
C	17	36.4	1.8	52297	4	US-08-705-551-1	Sequence 1, Appl	
C	18	36.4	1.8	72604	4	US-09-268-992-7	Sequence 1, Appl	
C	19	36.4	1.8	72604	4	US-09-657-474-7	Sequence 7, Appl	
C	20	36.2	1.8	236	4	US-09-453-702B-165	Sequence 7, Appl	
C	21	36.2	1.8	2372	4	US-09-145-391-1	Sequence 165, App	
C	22	36.2	1.8	3602	4	US-09-402-929-1	Sequence 1, Appl	
C	23	36	1.8	2007	3	US-08-747-221B-36	Sequence 1, Appl	
C	24	36	1.8	2007	3	US-08-747-221B-38	Sequence 36, Appl	
C	25	36	1.8	2007	4	US-09-005-051-36	Sequence 36, Appl	
C	26	36	1.8	2007	4	US-09-005-051-38	Sequence 38, Appl	
C	27	36	1.8	2700	3	US-09-315-861-1	Sequence 1, Appl	

28	3.6	1.8	2700	4	US-09-398-839A-43	Sequence 43, Appl1
29	36	1.8	5852	1	US-07-867-106-2	Sequence 2, Appl1
30	35.6	1.8	6702	3	US-08-987-439-2	Sequence 2, Appl1
31	35.6	1.7	817	4	US-08-885-456-1	Sequence 1, Appl1
32	35.6	1.7	817	4	US-09-625-918-1	Sequence 1, Appl1
33	35.6	1.7	1930	4	US-09-724-864-9	Sequence 9, Appl1
34	35.6	1.7	3812	4	US-09-784-834-9	Sequence 1, Appl1
35	35.6	1.7	3852	1	US-07-867-106-2	Sequence 1, Appl1
36	35.4	1.7	6611	1	US-08-280-443-1	Sequence 2, Appl1
37	35.4	1.7	6611	1	US-08-457-459-1	Sequence 1, Appl1
38	35.4	1.7	6611	1	US-08-555-678-1	Sequence 1, Appl1
39	35.4	1.7	6611	5	PCR-US95-02275-1	Sequence 1, Appl1
40	35.4	1.7	4581	1	US-08-524-751-1	Sequence 1, Appl1
41	35.2	1.7	1737	1	US-08-202-056-4	Sequence 4, Appl1
42	35.2	1.7	1737	1	US-08-076-093A-3	Sequence 3, Appl1
43	35.2	1.7	1737	1	US-08-701-265-3	Sequence 3, Appl1
44	35.2	1.7	1737	1	US-08-284-586-3	Sequence 3, Appl1
45	35.2	1.7	1737	2	US-08-805-478-3	Sequence 3, Appl1

## ALIGNMENTS

```

RESULT 1
US-09-134-001C-1334
: Sequence 1334, Application US/091344001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GNC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: CURRENT FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 1334
: LENGTH: 582
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
: US-09-134-001C-1334

```

Query Match

best local similarity	48.0%;	Pred. No. 5.2e-09;							
Matches	250;	Conservative	0;	Mismatches	265;	Indels	6;	Gaps	2

	Matches	250; Conservative	0; Mismatches	265; Indels	6; Gaps
QY	82	TCCTGTAGTGTCTTCGCGGGCCAGGGGACCGCGCAATCCACGCTGGCGTTCAGCTACG	14	14	
Db	24	TTTATGCTATTCTTAGTGTCCAACTGGTTCAGGTAACAGTAAGTATTTGAAGTTGCG	83		
QY	142	CCACGGCGCTGGCGGTAGATCGTCACGGCGTCCATCGAGGTCTATGAGAGCCCTAGA	20		
Db	84	TAAAAATTTAATGAGAAATATTATACGGAGATTCAATGCAAGTCTTATCAAGAAATGA	143		
QY	202	CATCATCACAACAAGGTTTTCTGCCAAGACAGACGAATCTGCCGCGACCAATGATCAG	261		
Db	144	TATGTGTACAGCAAAAGTTCAACTGAAGAATGAGATATACACATTATATGATAGA	203		
QY	262	CTTTGSGATCCCTCTTGTGACCAATTACACAGTGTGGACTTCGAAATTAGACCACTGC	321		
Db	204	TATTTTCCCTCCAGATGCTCTCTTTCTTGATAT - -GAATTTAAAAAAGGCGCAAAA	260		
QY	322	TCGTATGAGATATATTTCGCCGAGCAAAATCCATATGTTGGAGAGAACCAATTA	381		
Db	261	ATATATTAAAGATATTCTAGAAAGGCCAAGTGCCATTATATAGCAGAGAAACAGACT	320		
QY	382	TTACATTGAATCTCTGCTGTGAAAAGTTCTGTCAATACCAACGCCAGAGATGGGCAC	441		
Db	321	ATATATTCATCTCTTATACAACTATGCTTTTGAAGATGAATTCATATCTGAAGTTAA	380		

```
RESULT 15
US-09-835-232-6/c
; Sequence 6, Application US/09835232
; Patent No. US20020098489A1
; GENERAL INFORMATION:
; APPLICANT: Leder, Philip
; APPLICANT: Leder, Benjamin
; TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 00383/052002
; CURRENT APPLICATION NUMBER: US/09/835,232
; CURRENT FILING DATE: 2001-04-12
; PRIORITY FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 180216
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1) (180216)
; OTHER INFORMATION: n = A,T,C or G
US-09-835-232-6

Query Match      2.0%; Score 39.8; DB 10; Length 180216;
Best Local Similarity 53.5%; Pred. No. 22;
Matches 83; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1885 TGAGCACTGTGACATGAGCTTGAGACCAAGACTTTGAATTTGGAGCTGCTCAT 1944
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102428 TTAGAAACCACTGACACATATATGACTTTTCAGATATTTAAGTTATAGGCTCCT 102369
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1945 GTGTGAGTTATATACAGCTGCTCTTCTATGAGTTACAATCTATATTTTATTGAAG 2004
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102368 GCGTACCTTATTTAATCTTGATTTTGTCTGAGGTAGCTAATATTTTATCATACNA 102309
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 2005 TTTAATAAAGAAAAAATTTACAGAAAAA 2039
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102308 GATAAATTTTAAAGAAATTAATGGAAGTATGAA 102274
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: April 21, 2003, 18:47:42  
Job time : 370.611 secs

ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1394)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-301-410

Query Match 2.1%; Score 42.4; DB 10; Length 1401;  
Best Local Similarity 79.0%; Pred. No. 0.11;  
Matches 49; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1980 TTCAATCTATATTTTATGAAGTTAATAAGAAAAATTACAGAAAAAAA 2039  
DB 1295 TTATATAATATTTTGTGAACTTMAAAAAA 1354

QY 2040 AA 2041  
DB 1355 AA 1356

RESULT 12  
US-09-878-574-11368  
Sequence 11368, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
PRIOR FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 11368  
LENGTH: 257  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: 701064366H1  
US-09-878-574-11368

Query Match 2.0%; Score 41.4; DB 10; Length 257;  
Best Local Similarity 52.6%; Pred. No. 0.065;  
Matches 90; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 645 GTGGTCCCTTGGAGTCTGTAGATTCTTAACCTTGACCTTGCTTCATCTG 704  
DB 12 GTGCTCTCCTTGATGTTGGTCTTATTGAGCTACACCGGTACCGTGTGGTCCC 71  
QY 705 ACCAGCAGTTCTAGATGACCGCTTGGATAGAGGTGATGACATCTGCTGGGC 764  
DB 72 TTAAGGAGCTGTGAGTGGGGGTGGATATTCCTACAGTCAACAAGGTTGCTGGT 131  
QY 765 TCTTGGAGCACTAGAGATTTTCACAGCGCTTAATCAAGAAATGTTT 815  
DB 132 TTGATAAGGAGAAAGAGCTTGATCAGAGTTTCATCGAAATATGTTT 182

RESULT 13  
US-09-878-574-11381  
Sequence 11381, Application US/09878574  
Patent No. US20020110548A1  
GENERAL INFORMATION:  
APPLICANT: Byrum, Joseph R.  
APPLICANT: La Rosa, Thomas J.  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(15401)B  
CURRENT APPLICATION NUMBER: US/09/878,574  
CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535  
PRIOR FILING DATE: 1999-06-14  
NUMBER OF SEQ ID NOS: 15775  
SEQ ID NO 11381  
LENGTH: 263  
TYPE: DNA  
ORGANISM: Glycine max  
OTHER INFORMATION: Clone ID: 701064379H1  
US-09-878-574-11381

Query Match 2.0%; Score 41.4; DB 10; Length 263;  
Best Local Similarity 52.6%; Pred. No. 0.066;  
Matches 90; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 645 GTGGTCCCTTGGAGTCTGTAGATTCTTAACCTTGACCTTGCTTCATCTG 704  
DB 10 GTGCTCTCCTTGATGTTGGTCTTATTGAGCTACACCGGTACCGTGTGGTCCC 69

QY 705 ACCAGCAGTTCTAGATGACCGCTTGGATAGAGGTGATGACATCTGCTGGGC 764  
DB 70 TTAAGGAGCTGTGAGTGGGGGTGGATATTCCTACAGTCAACAAGGTTGCTGGT 129

QY 765 TCTTGGAGCACTAGAGATTTTCACAGCGCTTAATCAAGAAATGTTT 815  
DB 130 TTGATAAGGAGAAAGAGCTTGATCAGAGTTTCATCGAAATATGTTT 180

RESULT 14  
US-09-938-842A-1242  
Sequence 1242, Application US/09938842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kreps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
FILE REFERENCE: SCRIPI300-3  
CURRENT APPLICATION NUMBER: US/09/938,842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227,866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264,647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300,111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 1242  
LENGTH: 2526  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1242

Query Match 2.0%; Score 39.8; DB 9; Length 2526;  
Best Local Similarity 49.8%; Pred. No. 1;  
Matches 101; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 408 TTCTGTCAATPACCAACCCAGAGATGGGCACTGGAAGTATGACCGAAATGG 467  
DB 1211 TACTAGCAGATACAAACACCGCTCAAGATCAAGATTCAGAACTGAAAGATCTGG 1270  
QY 468 AGCTTGAAGAAGAGATGCTTCTTACTTCACAAAGCTTAAGCAGGTGACCCAGAAA 527  
DB 1271 ATTCTGGAAGGAACCGCAATCGAAAGTTAAACGAAAGACACAGAAAGCTTAAGACT 1330  
QY 528 TGGCTGCCAAGCTGCATCCATGACAAACGCAAGTGGCCAGGAGCTTCAAGTTTGG 587  
DB 1331 TGATTAACAAGTTGAATCCCATGAAACGTAATCGAGAGCAACAAGAGAGTTCTTG 1390  
QY 588 AAGAAACGAATCTCTCATAGT 610  
DB 1391 AAGCATCTGGAGTTGCTGATACT 1413

```

OY 262 CTTGTGGATCCTCTGTGACCAATTACACA 292
      | | | | | | | | | | | | | | | |
Db 34 TATCTTGATCCTGATGATCATTTTCGCA 4

RESULT 10
US-09-925-302-261
: Sequence 261, Application US/09925302
: Patent No. US20020044941A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA104
: CURRENT APPLICATION NUMBER: US/09/925,302
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05918
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 896
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 261
: LENGTH: 2952
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (20)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (33)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (35)
: OTHER INFORMATION: n equals a,t,g, or c
: US-09-925-302-261

Query Match 2.1%; Score 43.4; DB 10; Length 2952;
Best Local Similarity 54.8%; Pred. No. 0.096;
Matches 86; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

OY 1885 TGAGGAGCCTGTGACATAGCACTGGAAGCAAGACCTTGAATTTGCCGAGCTCTAT 1944
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2795 TGTTCATCTGTGTATTTTGTATTTAAAGATCAAAATTAATTGACATGCTGACATTTCT 2854
OY 1945 GTGTGAGTATTATTCACAGCTGCTCTTCTATTTAGATTACAATCTAATTTTATTGAAG 2004
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2655 ATACTCTGTTTAACACTGAGGATATCTCATTTTGCCCATGTAAATTTTCTTAATAAT 2914
OY 2005 TTTAATTAAGAAAAAATTTACAAGAAAAAATTTAAAA 2041
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2915 TGACAAAAACAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 2951

RESULT 11
US-09-925-301-410
: Sequence 410, Application US/09925301
: Patent No. US20020052308A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA106
: CURRENT APPLICATION NUMBER: US/09/925,301
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05882
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 1694
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 410
: LENGTH: 1401
: TYPE: DNA

```

Db 302 TGTGTACGTCACATCTTGTGTC 324

## RESULT 7

US-09-070-927A-106/c

Sequence 106, Application US/09070927A

Patent No. US20020120116A1

GENERAL INFORMATION:

APPLICANT: Charles A. Kunsch

Patrick J. Dillon

Steven Barash

TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

NUMBER OF SEQUENCES: 982

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/070,927A

FILING DATE: 04-May-2000

CLASSIFICATION: &lt;Unknown&gt;

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 60/046,655

FILING DATE: 1997-05-16

APPLICATION NUMBER: 60/044,031

FILING DATE: 1997-05-06

APPLICATION NUMBER: 60/066,009

FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:

LENGTH: 8519 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 106:

US-09-070-927A-106

Query Match

Best Local Similarity 2.5%; Score 50.4; DB 10; Length 8519;

Matches 160; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

Db 87 TAGTATCTCGGGCCAGCGGACCGCAATCCAGCTGCGCTTCAGCTAGCCAGC 146

Db 2425 TAGTATCTCGGGCCAGCGGACCGGTAAGCAATGAGTATGCTTACAAAA 2366

Db 147 GGCCTCGGGTGAAGTGTACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 206

Db 2365 AATTATGCTGTAATATATGAGCGGTGATGCTGCTGCTGCTGCTGCTGCTG 2306

Db 207 TCACCAACAGGTTTCTGCCAGAGAGAGAGATCTGCCGACACATGATGCTTGG 266

Db 2305 GCACCGTAAAGTACAGAAACAGAAAGAGAGATCTGCTGCTGCTGCTGCTGCTG 2246

Db 267 TGATCTCTCTGTCAGCAATATACAGTGTGAGCTTCAAGAAATAGAGCAACTGCTGTA 326

Db 2245 GTGAAGTACGAAACATATT--CTGCCGCTGATTTTCAGAAAGAGGGTGTCAAAAA 2189

Db 327 TTGAAGATATATTTTGGCCGAGACAAATTCCTATTTGTTGGAGAGAACATATTATTA 386

Db 2188 TCAAGAAATTTACTGAAAAAGGAAACTACCAATATTCTGTGAGAGAACAGTTGTATA 2129

Db 387 TTGAATCTGCTGCTGGAAGTTC 410

Db 2128 TTCAATCAGCTCTTTATGATTTTC 2105

## RESULT 8

US-08-781-986A-800/c

Sequence 800, Application US/08781986A

Publication No. US20030054436A1

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5255

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/781,986A

FILING DATE:

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Benson, Bob

REGISTRATION NUMBER: 30,446

REFERENCE/DOCKET NUMBER: PB248PP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 800:

SEQUENCE CHARACTERISTICS:

LENGTH: 989 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 800:

US-08-781-986A-800

Query Match

Best Local Similarity 2.4%; Score 49.8; DB 7; Length 989;

Matches 136; Conservative 0; Mismatches 123; Indels 3; Gaps 1;

Db 144 AGCGGCTCGGGGTGAATGCTACAGCGCTGCTGCTGCTGCTGCTGCTGCTG 203

Db 986 AGCTATCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 927

Db 204 TCATCACCACCAAGGTTTCTGCCAGAGAGAGAGATCTGCCGACACATGATGAGCT 263

Db 926 TTGGAATGCAAAAGTACACCTGAGAAATGATGATTCACATCATTTAATTTGATA 867

Db 264 TTGGAATGCTCTGTCAGCAATATACAGTGTGAGCTTGCAGAAATAGAGCAACTGCTC 323

Db 866 TCTTGATCTCGATGATGATGATTTT--CAGATATGATTTAAGCAATGACAGAAAT 810

Db 324 TGATGAAGATATATTTTGGCCGAGACAAATTCCTATTTGTTGGAGAGAACATTTATT 383

Db 809 TAATTTCTGATTAACCAATAGAGGTAAAGTTTCAATCATGAGAGTGAACAGCTTAT 750

Db 384 ACATGAATCTCTGCTGAGAA 405

Db 749 ATATTCAATCATTTAATATATTA 728

```

RESULT 5
US-09-864-761-26692/C
Sequence 26692, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmca-x-1
CURRENT APPLICATION NUMBER: US/09/864,761
PCT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 26692
LENGTH: 138
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL033527.25
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HELLA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: AF074918.1, EVALUATE 3.00e-72
OTHER INFORMATION: EST_HUMAN HIT: BE242161.1, EVALUATE 3.00e-72
OTHER INFORMATION: SWISSPROT HIT: O9KAC3, EVALUATE 2.00e-03
US-09-864-761-26692
Query Match 6.8%; Score 138; DB 10; Length 138;
Best Local Similarity 100.0%; Fred.No. 1e-30; Mismatches 0; Indels 0; Gaps 0;
Matches 138; Conservative 0;

```

QY	185	GTCTTGAAGGCTTGACATCATCAACAAGAGTTTCTGCCAAGAGAGAGATCTGC	244
DB	138	GTCTTGAAGGCTTGACATCATCAACAAGAGTTTCTGCCAAGAGAGAGATCTGC	79
QY	245	CGGCACCAATGATGAGCTTTGTGATCCTCTTGTGACCAATTAACAGTGGTGACTTC	304
DB	78	CGGCACCAATGATGAGCTTTGTGATCCTCTTGTGACCAATTAACAGTGGTGACTTC	19
QY	305	AGAAATAGAGCACTGCT 322	
DB	18	AGAAATAGAGCACTGCT 1	
RESULT 6			
	US-09-738-626-2133		
	Sequence 2133, Application US/09738626		
	Publication No. US20020197605A1		
	GENERAL INFORMATION:		
	APPLICANT: NAKAGAWA, SATOSHI		
	APPLICANT: MIZOGUCHI, HIROSHI		
	APPLICANT: ANDO, SEIKO		
	APPLICANT: HAYASHI, MIKIRO		
	APPLICANT: OCHIAI, KEIKO		
	APPLICANT: YOKOI, HARUHIKO		
	APPLICANT: TATEISHI, NAOKO		
	APPLICANT: SENO, AKIHIRO		
	APPLICANT: IKEDA, MASATO		
	APPLICANT: OKAZAKI, AKIO		
	TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES		
	FILE REFERENCE: 249-125		
	CURRENT APPLICATION NUMBER: US/09/738, 626		
	CURRENT FILING DATE: 2000-12-18		
	PRIOR APPLICATION NUMBER: JP 99/377484		
	PRIOR FILING DATE: 1999-12-16		
	PRIOR APPLICATION NUMBER: JP 00/159162		
	PRIOR FILING DATE: 2000-04-07		
	PRIOR APPLICATION NUMBER: JP 00/280988		
	PRIOR FILING DATE: 2000-08-03		
	NUMBER OF SEQ ID NOS: 7059		
	SOFTWARE: Patencin ver. 3.0		
	SEQ ID NO 2133		
	LENGTH: 903		
	TYPE: DNA		
	ORGANISM: Corynebacterium glutamicum		
	US-09-738-626-2133		
	Query Match	3.7%;	Score 75; DB 9; Length 903;
	Best Local Similarity	54.2%;	Pred. No. 1.8e-11;
	Matches 175; Conservative	0;	Mismatches 145; Indels 3; Gaps 1;
QY	78	TACCTCTTGTATGATGATCTCGGGGCGACCGGCAACCGGCAATTCACAGCTGGCGTTGCAC	137
DB	5	TACACCGATCCAGTGGTGTGAGACCCACATCTGGAAAAATCAGTTGGAAATGCTC	64
QY	138	TAGGCGAGGCGTGGCGGTGAGATGCTGAGCGTGTACTGCATGACAGGTCTATGAAGGCC	197
DB	65	TAGCCACAGCTTGTGACGTTGAAGTAGTAATGTGATTCATGCAAGCTGTACAAAGCA	124
QY	198	TAGACATCATCAACAAGGTTTCTGCCCAAGAGAGAGAGAAATCTGCGGACCAACATGA	257
DB	125	TGACATCGGCGACGGAAGGTGACTGTGGAAGACGCAAGGCAATTCGCGCATCAGC	184
QY	258	TGAGCTTGTGATCTCTTGTGACCAATTAACAGTGTGAGACTTCAGAAATAGAGCA	317
DB	185	TGAGTGTGAGGACCTTACCGAAC--TGGCTGAGTGGCGAGTTTCATTCGAGCGC	241
QY	318	GTGCTCTGATTAAGATATATTTGCCGAGACAAATTCCTATTGTTGTGGAGAGACCA	377
DB	242	TTCGCGAGTGAAGATATATATGTGTCGCGGCAAAACCCCATCTTGTGTGGCGCTCA	301
QY	378	ATTATTACATTGATCTGCTC 400	

```
;; CURRENT FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO: 4310
;; LENGTH: 318
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-796-692-4310
```

Query Match 15.6%; Score 318; DB 9; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9e-84;

Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1690 AAGAAGTTTATGTCCTGAGTCTGGCTAAATATCTAATTCAGATGCTTTGTAG 1749
    |||
DB 1 AAAGAAGTTTATGTCCTGAGTCTGGCTAAATATCTAATTCAGATGCTTTGTAG 60
QY 1750 ATGACTGAGTATTTGTAGCCACATATTGGAGTTCTGATTTGAGTGAAGCGAGAA 1809
    |||
DB 61 ATGACTGAGTATTTGTAGCCACATATTGGAGTTCTGATTTGAGTGAAGCGAGAA 120
QY 1810 AGGGCATCTCCATTCAGATGATTAAGTACCAACTAGTTCTGGAAATTCACAGAGA 1869
    |||
DB 121 AGGGCATCTCCATTCAGATGATTAAGTACCAACTAGTTCTGGAAATTCACAGAGA 180
QY 1870 AGGAGGAATCAGACTGAGGAAGCTGTGACATAGAGACTTGAAGACCTTTGAAT 1929
    |||
DB 181 AGGAGGAATCAGACTGAGGAAGCTGTGACATAGAGACTTGAAGACCTTTGAAT 240
QY 1930 TTGGCAGCTGCTCATGTGAGATTATTCACACTGCTGTTCTTATTCAGTTACAATCT 1989
    |||
DB 241 TTGGCAGCTGCTCATGTGAGATTATTCACACTGCTGTTCTTATTCAGTTACAATCT 300
QY 1990 ATATTTTATTCAGTTT 2007
    |||
DB 301 ATATTTTATTCAGTTT 318
```

RESULT 4  
US-09-864-761-10050/c

; Sequence 10050, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; GENE EXPRESSION ANALYSIS BY MICROARRAY

```
;; FILE REFERENCE: Aeonica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
;; SEQ ID NO: 10050
;; LENGTH: 457
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL033527.25
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.98
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-10050
```

Query Match 7.1%; Score 145; DB 10; Length 457;  
Best Local Similarity 100.0%; Pred. No. 2e-32;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 183 AGGTCTATGAGGCTCAGATCATCATCACCACCAAGTTTCTGGCCAAAGACAGAAATCT 242
    |||
DB 450 AGGTCTATGAGGCTCAGATCATCATCACCACCAAGTTTCTGGCCAAAGACAGAAATCT 391
QY 243 GCCGGCACCACATGATCAGCTTTGTGATCTCTTGTGACCAATTCACAGTGTGAGT 302
    |||
DB 390 GCCGGCACCACATGATCAGCTTTGTGATCTCTTGTGACCAATTCACAGTGTGAGT 331
QY 303 TCAGAAATGAGCACTGCTCTGAT 327
    |||
DB 330 TCAGAAATGAGCACTGCTCTGAT 306
```



```

OY 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
    |||||||
DB 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
OY 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
    |||||||
DB 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
OY 181 GCAGCTCTATAGAGCCCTAGACATCATCATCAACAAGGTTTCTGCCCAAGACAGAGANT 240
    |||||||
DB 181 GCAGCTCTATAGAGCCCTAGACATCATCATCAACAAGGTTTCTGCCCAAGACAGAGANT 240
OY 241 CTGCGGACACACATGATGATCTTTGGATGCTCTTGACCAATTCACAGTGTGGA 300
    |||||||
DB 241 CTGCGGACACACATGATGATCTTTGGATGCTCTTGACCAATTCACAGTGTGGA 300
OY 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
    |||||||
DB 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
OY 361 TGTGTGGAGAGAACCAATTTATACATGAACTCTGCTGGAAGTTCTTGTCAATAC 420
    |||||||
DB 361 TGTGTGGAGAGAACCAATTTATACATGAACTCTGCTGGAAGTTCTTGTCAATAC 420
OY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAGA 480
    |||||||
DB 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAGA 480
OY 481 GGATGCTCTTACTTTCACAAACGGCTTAAGCAGGTGAGCCAGAAATGCTGCCAAGCT 540
    |||||||
DB 481 GGATGCTCTTACTTTCACAAACGGCTTAAGCAGGTGAGCCAGAAATGCTGCCAAGCT 540
OY 541 GCATCCACATGACAAAGCCAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGANT 600
    |||||||
DB 541 GCATCCACATGACAAAGCCAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGANT 600
OY 601 CTCATCATGATGATTTCTCCATCGTCACACATA-CGGAAGAAAGTGGTGTGCCCC-TTGA 658
    |||||||
DB 601 CTCATCATGATGATTTCTCCATCGTCACACATA-CGGAAGAAAGTGGTGTGCCCC-TTGA 658
OY 659 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 694
    |||||||
DB 659 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 694
OY 661 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 696
    |||||||
DB 661 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 696

RESULT 2
; Sequence 364, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 364
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (619)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (625)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (655)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

; NAME/KEY: SITE
; LOCATION: (673)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (713)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-364

Query Match      32.4%; Score 661.8; DB 10; Length 717;
Best Local Similarity 98.4%; Pred. No. 2.4e-185;
Matches 685; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

OY 1 CTGCCATAGATGAGCGCTCGGCGGCTGCAAGACAGTCTCTGAGGAGTGGGCTGAG 60
    |||||||
DB 1 CTGCCATAGATGAGCGCTCGGCGGCTGCAAGACAGTCTCTGAGGAGTGGGCTGAG 60
OY 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
    |||||||
DB 61 GGGCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGGGCCAGGGGACCGGCAATTC 120
OY 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
    |||||||
DB 121 CACGCTGGCGTTGAGCTAGAGCCAGCGGCTCGGGGTGAGATCGTCACGCTGACTCCAT 180
OY 181 GCAGCTCTATAGAGCCCTAGACATCATCATCAACAAGGTTTCTGCCCAAGACAGAGANT 240
    |||||||
DB 181 GCAGCTCTATAGAGCCCTAGACATCATCATCAACAAGGTTTCTGCCCAAGACAGAGANT 240
OY 241 CTGCGGACACACATGATGATCTTTGGATGCTCTTGACCAATTCACAGTGTGGA 300
    |||||||
DB 241 CTGCGGACACACATGATGATCTTTGGATGCTCTTGACCAATTCACAGTGTGGA 300
OY 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
    |||||||
DB 301 CTTCAGAAATAGAGCACTGCTGTGATTAAGATATATTTGCCGAGACAAATTCCTAT 360
OY 361 TGTGTGGAGAGAACCAATTTATACATGAACTCTGCTGGAAGTTCTTGTCAATAC 420
    |||||||
DB 361 TGTGTGGAGAGAACCAATTTATACATGAACTCTGCTGGAAGTTCTTGTCAATAC 420
OY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAGA 480
    |||||||
DB 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTGAAGAAGA 480
OY 481 GGATGCTCTTACTTTCACAAACGGCTTAAGCAGGTGAGCCAGAAATGCTGCCAAGCT 540
    |||||||
DB 481 GGATGCTCTTACTTTCACAAACGGCTTAAGCAGGTGAGCCAGAAATGCTGCCAAGCT 540
OY 541 GCATCCACATGACAAAGCCAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGANT 600
    |||||||
DB 541 GCATCCACATGACAAAGCCAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGANT 600
OY 601 CTCATCATGATGATTTCTCCATCGTCACACATA-CGGAAGAAAGTGGTGTGCCCC-TTGA 658
    |||||||
DB 601 CTCATCATGATGATTTCTCCATCGTCACACATA-CGGAAGAAAGTGGTGTGCCCC-TTGA 658
OY 659 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 694
    |||||||
DB 659 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 694
OY 661 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 696
    |||||||
DB 661 GGTCTCTGAAGTCTCTTAACCCCTTGACCTTTGG 696

RESULT 3
; Sequence 4310, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THE
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:55:39 ; Search time 235.611 Seconds  
(without alignments)  
8702.120 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041

Sequence: 1 CTGCCATAAGATGGCGTCGCG.....TTTACAGAGAAAAA 2041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	663	32.5	717	10	US-09-764-853-166 Sequence 166, App
2	661.8	32.4	717	10	US-09-764-853-364 Sequence 364, App
3	318	15.6	318	9	US-09-796-692-4310 Sequence 4310, App
4	145	7.1	457	10	US-09-864-761-10050 Sequence 10050, A
5	138	6.8	138	10	US-09-864-761-26692 Sequence 26692, A
6	75	3.7	903	9	US-09-738-626-2133 Sequence 2133, App
7	50.4	2.5	8519	10	US-09-070-927A-106 Sequence 106, App
8	49.8	2.4	989	7	US-08-781-986A-800 Sequence 800, App
9	47.8	2.3	846	10	US-09-939-980-65 Sequence 65, App1
10	43.4	2.1	2952	10	US-09-925-302-261 Sequence 261, App
11	42.4	2.1	1401	10	US-09-925-301-410 Sequence 410, App
12	41.4	2.0	257	10	US-09-878-574-11368 Sequence 11368, A
13	41.4	2.0	263	10	US-09-878-574-11381 Sequence 11381, A
14	39.8	2.0	2526	9	US-09-938-842A-1242 Sequence 1242, App
15	39.8	2.0	180216	10	US-09-835-232-6 Sequence 6, App1
16	39.4	1.9	368	10	US-09-960-352-5552 Sequence 5552, App
17	39.2	1.9	632	9	US-09-992-331-4 Sequence 4, App1
18	39.2	1.9	3092	10	US-09-925-302-310 Sequence 310, App
19	38.8	1.9	377	10	US-09-960-352-7419 Sequence 7419, App

C	20	38.8	1.9	424	10	US-09-960-352-11218	Sequence 11218, A
C	21	38.8	1.9	722	12	US-10-001-878-104	Sequence 104, App
C	22	38.4	1.9	265	10	US-09-969-373-1074	Sequence 1074, App
C	23	38.4	1.9	454	10	US-09-960-352-8765	Sequence 8765, App
C	24	38.4	1.9	486	9	US-09-832-292-36	Sequence 36, App1
C	25	38.2	1.9	1170	10	US-09-729-674-119	Sequence 119, App
C	26	38	1.9	1605	9	US-09-809-391-187	Sequence 187, App
C	27	38	1.9	2821	9	US-09-736-457-1669	Sequence 1669, App
C	28	38	1.9	2821	9	US-09-902-941-1669	Sequence 1669, App
C	29	38	1.9	2821	9	US-09-849-626-1669	Sequence 1669, App
C	30	38	1.9	2821	9	US-10-017-754-1669	Sequence 1669, App
C	31	37.8	1.9	933	9	US-10-260-877-31	Sequence 31, App1
C	32	37.8	1.9	980	10	US-09-452-229-21	Sequence 21, App1
C	33	37.8	1.9	1231	10	US-09-891-126-3	Sequence 3, App1
C	34	37.8	1.9	4006	10	US-09-925-300-580	Sequence 580, App
C	35	37.8	1.9	16472	9	US-10-079-854-199	Sequence 199, App
C	36	37.8	1.9	16472	10	US-09-764-878-199	Sequence 199, App
C	37	37.8	1.9	32249	9	US-10-079-854-202	Sequence 202, App
C	38	37.8	1.9	32249	10	US-09-764-878-202	Sequence 202, App
C	39	37.4	1.8	213	10	US-09-919-580-736	Sequence 736, App
C	40	37.4	1.8	349	10	US-09-770-791-712	Sequence 712, App
C	41	37.4	1.8	373	10	US-09-960-352-836	Sequence 836, App
C	42	37.4	1.8	2025	9	US-09-809-391-316	Sequence 316, App
C	43	37.4	1.8	2346	9	US-09-809-391-193	Sequence 193, App
C	44	37.4	1.8	2972	9	US-10-057-832-1	Sequence 1, App1
C	45	37.2	1.8	430	10	US-09-960-352-2920	Sequence 2920, App

## ALIGNMENTS

RESULT 1  
US-09-764-853-166  
Sequence 166, Application US/09764853  
Patent No. US20020096072A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PJ206  
CURRENT APPLICATION NUMBER: US/09/764,853  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 939  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 166  
LENGTH: 717  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (619)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (625)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (655)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (673)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (713)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-853-166

Query Match 32.5%; Score 663; DB 10; Length 717;  
Best Local Similarity 98.4%; Pred. No. 1.1e-185;  
Matches 665; Conservative 3; Mismatches 6; Indels 2; Gaps 2;

QY 1 CTGCCATAAGATGGCGTCGCGGCTGCACGACGAGTCCTCTGTGGGCACTGGCTCAG 60  
DB 1 CTGCCATAAGATGGCGTCGCGGCTGCACGACGAGTCCTCTGTGGGCACTGGCTCAG 60



```

DB 1325 AGAAGAGGTGGAGCTTAGACGCTGTGAGTCCAGAGAGTCAAGTAATTCCTCCAGAC 1384
QY 1268 TATAACAAAGACCTTAAGGGAAGGATCCCAAGGAGATGATCAAGAGCTCAAAATGC 1327
DB 1385 TGTGACCCGAGCATCGAGGGGGAATCTCCGGGGAGACATCAATCAAGAGCTGAAGGCC 1444
QY 1328 AGGCTTAGAGACATGTCAGTGGCTTTGGAAGAGTGGTGGGATCCAGTTGAGAGG 1387
DB 1445 AGGTGTTGAGAGACATGCTAGTGGCTTTGAGAG - GACGTGGGATCAAGCCAGAGG 1503
QY 1388 GAGGGGTATGTTGTCTCCA - CTCGGGCAAGAGAGTGTATGCGAATTCCTGCAAT 1446
DB 1504 GAGGGAGTGTGCTCTCCAGCGCTGAGCTGAGATGCTGGAG - ----- 1550
QY 1447 GCAGAAAAGCTCCACCATTTTCTTTGATGTTGATTTAAAGCTACCTCTCTATPAT 1506
DB 1551 -----AAGGCCCAACATCTTCTTTCATTCGTGCT-----GTGCTCTGAGT 1593
QY 1507 AGAAGACAGAGTCTTGTGAGCTCCTTGTGTGCTGATGCTGGAATGATGATGCTTC 1566
DB 1594 GGAACACAGAGCCCTT - TCAGCTCCTTGTGCTGCTGTGCTGTGATGATGATGATTC 1652
QY 1567 AGGAAGCATTTTCTTTCTTTGAACTTAAGCTTCTATATTAAGAGCAGACAGAT 1626
DB 1653 AG---AGTGGGATTTTCTTTGAACTTAAGCTTCTATATTAAGAGCAGACAGAT 1709
QY 1627 TCACATTTTATACATGAGATCTTCTTTGGTGAATPACAGATGATGATGCTTCCT 1686
DB 1710 CGCACATTTTCTTCTGAGATCTTTTATGTTGATGATGATGATGATGATGATGATGAT 1768
QY 1687 TTAAGAGATTTATGCTCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
DB 1769 TTAAGAGATTTATGCTCCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1828
QY 1747 TAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1806
DB 1829 TAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1888
QY 1807 GAAAGGCGCATCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1866
DB 1889 GAGTAG-----TGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1938
QY 1867 AGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1925
DB 1939 CAGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1998
QY 1926 AATTTGCGAGCTGCTGATG 1947
DB 1999 AGTCTGCTGAGCTGCTGCTG 2020

```

RESULT 14  
US-09-532-315B-3472

```

; Sequence 3472, Application US/09532315B
; GENERAL INFORMATION:
; APPLICANT: Selhamer, Jeffrey J.
; APPLICANT: Deleane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
; FILE REFERENCE: PD-1002 CIP
; CURRENT APPLICATION NUMBER: US/09/532,315B
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 07/916,491
; PRIOR FILING DATE: 1992-07-17
; PRIOR APPLICATION NUMBER: 07/977,780
; PRIOR FILING DATE: 1992-11-19
; PRIOR APPLICATION NUMBER: 08/100,523
; PRIOR FILING DATE: 1993-08-03
; PRIOR APPLICATION NUMBER: 09/008,119

```

```

; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 08/196,364
; PRIOR FILING DATE: 1994-02-14
; PRIOR APPLICATION NUMBER: 08/282,991
; PRIOR FILING DATE: 1994-07-28
; PRIOR APPLICATION NUMBER: 08/438,571
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: 08/179,873
; PRIOR FILING DATE: 1994-01-11
; PRIOR APPLICATION NUMBER: 08/504,732
; PRIOR FILING DATE: 1995-07-20
; PRIOR APPLICATION NUMBER: 08/137,951
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42212
; SOFTWARE: PERL Program
; SEQ ID NO: 3472
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Inycle ID No: hu01286535
US-09-532-315B-3472

```

```

Query Match      26.5%      Score 540.8; DB 6; Length 555;
Best Local Similarity 99.5%; Pred. No. 1.6e-132;
Matches 553; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1216 ATTGACTCAGATGCTGTCAACACATPAGAAAGTGAAGTGTGTTCCAGACTATACAA 1275
DB 1 ATTGACTCAGATGCTGTCAACACATPAGAAAGTGAAGTGTGTTCCAGACTATACAA 60
QY 1276 AGAAGCTTAAGGAGGATCCCAAGGAGCAGATGATCAAGAGCTGAAGTGAAGTGAAGTGA 1335
DB 61 AGAAGCTTAAGGAGGATCCCAAGGAGCAGATGATCAAGAGCTGAAGTGAAGTGAAGTGA 120
QY 1336 AGAGACATGCTCAGTGGGCTTTGGAAGGCTGTGGGATCCAGTTCAGAGAGGAGGTA 1395
DB 121 AGAGACATGCTCAGTGGGCTTTGGAAGGCTGTGGGATCCAGTTCAGAGAGGAGGTA 180
QY 1396 TGTGTTGCTCCAGCTGCGCAAGAGAGTGTATGCGGAATTCCTGATGAGCAAAAG 1455
DB 181 TGTGTTGCTCCAGCTGCGCAAGAGAGTGTATGCGGAATTCCTGATGAGCAAAAG 240
QY 1456 CTCACCATTTTCTTTGATGTTGTTTAAAGTCTCAGCTTCTATATAGAAACAGC 1515
DB 241 CTCACCATTTTCTTTGATGTTGTTTAAAGTCTCAGCTTCTATATAGAAACAGC 300
QY 1516 AGGCTGTTGAGCTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1575
DB 301 AGGCTGTTGAGCTCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 360
QY 1576 TTTTCTTTTCTTTGAACTTAAGGTTCTATATTAATTAAGGAGCAGAGTTCACATTT 1635
DB 361 TTTTCTTTTCTTTGAACTTAAGGTTCTATATTAATTAAGGAGCAGAGTTCACATTT 420
QY 1636 TATACATAGAGATCTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1695
DB 421 TATACATAGAGATCTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1755
QY 1696 GTTTATGTCCTGACTGCTGCTAAATATATATATATATATATATATATATATATATAT 1755
DB 480 GTTTATGTCCTGACTGCTGCTAAATATATATATATATATATATATATATATATATAT 539
QY 1756 GAAGTATTTGAGGCC 1771
DB 540 GAAGTATTTGAGGCC 555

```

RESULT 15  
US-09-532-315B-3471  
; Sequence 3471, Application US/09532315B





Db	1262	AAAGCATTTTTTTTCTTTGAACCTTAAGGCTTCAATTATTAAGACACACAGATTC	1321
Qy	1630	ACATTTTATACATGAGGATCTTCTTTGTGTGAATACAGGATTCAGTCATCCCTT	1689
Db	1322	ACATTTTATACATGAGGATCTTCTTTGTGTGAATACAGGATTCAGTCATCCCTT	1380
Qy	1690	AAGAAGTTTATGTCCTGAGCTCTGGCTAAATTTATTAATTTCCAGATGCTTTGTAG	1749
Db	1381	AAAGAAGTTTATGTCCTGAGCTCTGGCTAAATTTATTAATTTCCAGATGCTTTGTAG	1440
Qy	1750	ATGACTGAAGTATTTGTGTGAGCCACATATTTGGAGTTCAGATTTGAGTGAATGGCAGAA	1809
Db	1441	ATGACTGAAGTATTTGTGTGAGCCACATATTTGGAGTTCAGATTTGAGTGAATGGCAGAA	1500
Qy	1810	AGGSCATCTTCATTGAGATGATTAAGTGACCAACCATAGTTCTCGGAATTTACACAGA	1869
Db	1501	AGGGCCATCTTCATTGAGATGATTAAGTGACCAACCATAGTTCTCGGAATTTACACAGA	1560
Qy	1870	AGGAGGGAATCAGACTGAGCAGAGCTGTGACATTCAGACTTGAACCAAGACATTGGAAT	1929
Db	1561	AGGAGGGAATCAGACTGAGCAGAGCTGTGACATTCAGACTTGAACCAAGACATTGGAAT	1620
Qy	1930	TTGCGAGCTGCTCATGTGTGAGTTATTAATCAGCTGTCTTCTATTGAGTACCAAAATCT	1989
Db	1621	TTGCGAGCTGCTCATGTGTGAGTTATTAATCAGCTGTCTTCTATTGAGTACCAAAATCT	1680
Qy	1990	ATAATTTTATGAAGTTAAATTAAGAAAAAATTTTCAA	2028
Db	1681	ATAATTTTATGAAGTTAAATTAAGAAAAAATTTTCAA	1719

```

RESULT 10
US-10-170-235-41124
: Sequence 41124, Application US/10170235
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig
: TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
: FILE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
: FILE REFERENCE: CLO01380
: CURRENT APPLICATION NUMBER: US/10/170,235
: CURRENT FILING DATE: 2003-03-17
: NUMBER OF SEQ ID NOS: 42514
: SEQ ID NO 41124
: LENGTH: 1271
: TYPE: DNA
: ORGANISM: HUMAN
: US-10-170-235-41124

```

Query Match	52.6%;	Score 1073.8;	DB 8;	Length 1271;
Best Local Similarity	98.4%;	Pred. No. 2.1e-273;		
Matches 1095; Conservative	0;	Mismatches 17;	Indels 1;	Gaps 1;

OY	916	TAGTACACGCTTAAAGAAAGAGACTGTGCCATTGCCCCCTGTATAGCGTTAGA	975
Db	160	TGCGTACGCGTGACTCCATCTCAGAGACCTGGTCCCATTTGCCCTGTATAGCGTTAGA	219
OY	976	GGATCTGATGTCTCGAAGTGGGAGAGATCTGTTCTTGAACCTGCTCTGGAATGTGTCA	10353
Db	220	GGATCTCGATGTCTCGAAGTGGGAGAGATCTGTTCTTGAACCTGCTCTGGAATGTGTCA	279
OY	1036	AAGTTATCCAGGGGCCACAGAGCTCTGACGACTCCCAATTAAGATGCCATTCATGAAGC	10955
Db	280	AAGTTATCTCAGGGGCCACAGAGCTCTGACGACTCCCAATTAAGATGCCATTCATGAAGC	339
OY	1096	TGAGACAAGAGAAGTTATACCTGTGTGACCTCTGTGTGATCGAATCATATTGGGGATCG	11555
Db	340	TGAGACAAGAGAAGTTATACCTGTGTGACCTCTGTGTGATCGAATCATATTGGGGATCG	399
OY	1156	CGAATGGGACGGCGACATTAATAATCCAAATCCCACTTGAAACCAACGAGAAAAGAAAG	12115
Db	400	CGAATGGGACGGCGACATTAATAATCCAAATCCCACTTGAAACCAACGAGAAAAGAAAG	459
OY	1216	ATTGGACTCAGATCTGTCTCAACACCATAGAAGATGAGATGTTTCCCAGACTATACAA	1275

```

Db      460  ATTGGACTAGATGCTGTCTACACCAATAGAAAGTCAGAGTGTTCCTCCAGACCTAAACAA  519
QY      1276  AGAACCCTAAAGGAGAGGAGATCCCCAGAGGCGAATATGATCAAGAGCTGAATATGCACCGTTTA  1335
Db      520  AGAACCCTAAAGGAGAGGAGATCCCCAGAGGCGAATATGATCAAGAGCTGAATATGCACCGTTTA  579
QY      1336  AGAGACATGTCAGTGGCCTTTTGAAAAGTGCTGGGGATCCAGTTCAAGAGGAGGGGCTA  1395
Db      580  AGAGACATGTCAGTGGCCTTTTGAAAAGTGCTGGGGATCCAGTTCAAGAGGAGGGGCTA  639
QY      1396  TGTTCCTCCAGTGTGGGCAAGAGAGTCTATGCGGATTTCTCTCTCATAGCAAGAAAG  1455
Db      640  TGTTCCTCCAGTGTGGGCAAGAGAGTCTATGCGGATTTCTCTCTCATAGCAAGAAAG  699
QY      1456  CTCGCCACATTTCTTTTGAATGATGGTTTTTAAGCTCAGCTTCCTATTAATAGAAACAGC  1515
Db      700  CTCGCCACATTTCTTTTGAATGATGGTTTTTAAGCTCAGCTTCCTATTAATAGAAACAGC  759
QY      1516  AGGCTCTGTACCTCCTTGTGTGGCTGATGTGTCTGGAATGATGTAAGTTCCAGGAAAGCA  1575
Db      760  AGGCTCTGTACCTCCTTGTGTGGCTGATGTGTCTGGAATGATGTAAGTTCCAGGAAAGCA  819
QY      1576  TTTTTCCTTTTCTTTGACCTTAAAGGTTCTATTATTATTAAGCACACAGATTCACACTTT  1635
Db      820  TTTTTCCTTTTCTTTGACCTTAAAGGTTCTATTATTATTAAGCACACAGATTCACACTTT  879
QY      1636  TATATCATAGAGATCTCTTGTGTGGTGAATACCAAGATGACATGCAATCCCTTTAAAGAA  1695
Db      880  TATATCATAGAGATCTCTTGTGTGGTGAATACCAAGATGACATGCAATCCCTTTAAAGAA  938
QY      1696  GTTTATATGCCCTGACTCTGGCTAAATATATCTAATTTCCAGATGCTTTGTAGATGACT  1755
Db      939  GTTTATATGCCCTGACTCTGGCTAAATATATCTAATTTCCAGATGCTTTGTAGATGACT  998
QY      1756  GAAGTATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTGAGTGCAGAGAAAGGCC  1815
Db      999  GAAGTATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTGAGTGCAGAGAAAGGCC  1058
QY      1816  ATCTCATGTGAGATGATTTAAGTGAACCAACATGTTCTGGAATTTACAGAGAGAGAGG  1875
Db      1058  ATCTCATGTGAGATGATTTAAGTGAACCAACATGTTCTGGAATTTACAGAGAGAGAGG  1118
QY      1876  GATCAGACTGAGGAAGCTGTGACATAGACATTGAAGACCAAGACACTTGAATTTGCGA  1935
Db      1119  GATCAGACTGAGGAAGCTGTGACATAGACATTGAAGACCAAGACACTTGAATTTGCGA  1178
QY      1936  GCTGCTCATGTGTGAGTTATATCACTGCTGCTTTCTATATTGAGTTCAAAATCTATATTT  1995
Db      1179  GCTGCTCATGTGTGAGTTATATCACTGCTGCTTTCTATATTGAGTTCAAAATCTATATTT  1238
QY      1996  TTAATGAAGTTAAATAAGAAATAATTTCAA  2028
Db      1239  TTAATGAAGTTAAATAAGAAATAATTTCAA  1271

RESULT 11
US-60-453-135-276
; Sequence 276, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

QY 1330 CGTTAAGACATGTCAGTGGCCCTTTGAAAGTGCTGGGATCCAGTTCAGAGGA 1389
DB 1022 CGTTAAGACATGTCAGTGGCCCTTTGAAAGTGCTGGGATCCAGTTCAGAGGA 1081
QY 1390 GGGGATGTTTGTCTCCAGTCTGGGCAAGAGTCTATGCGGAATCTCTGCAATGA 1449
DB 1082 GGGGATGTTTGTCTCCAGTCTGGGCAAGAGTCTATGCGGAATCTCTGCAATGA 1141
QY 1450 GAAAGCCTCCACCATTTCTTTGATGTGCTTTAAAGTCTCAGTCTCTTAATAGA 1509
DB 1142 GAAAGCCTCCACCATTTCTTTGATGTGCTTTAAAGTCTCAGTCTCTTAATAGA 1201
QY 1510 AACAGAGCTCTGTGAGCTCTGTGCTGATGTCTGGAATGATGATGATGATGATG 1569
DB 1202 AACAGAGCTCTGTGAGCTCTGTGCTGATGTCTGGAATGATGATGATGATGATG 1261
QY 1570 AAAGCATTTTCTTTCTTTGAACTTAAAGTCTTATTAATTAAGACAGACAGATTC 1629
DB 1262 AAAGCATTTTCTTTCTTTGAACTTAAAGTCTTATTAATTAAGACAGACAGATTC 1321
QY 1630 ACATTTTATACATGAGATCTTCTTGTGTGAATACAGAGATGATGATGATGATGATG 1689
DB 1322 ACATTTTATACATGAGATCTTCTTGTGTGAATACAGAGATGATGATGATGATGATG 1380
QY 1690 AAAGAGTTTATGCTCTGAGTCTGCTAAATTAATTAATTAATTAATTAATTAATTA 1749
DB 1381 AAAGAGTTTATGCTCTGAGTCTGCTAAATTAATTAATTAATTAATTAATTAATTA 1440
QY 1750 ATGACTGAAATTTTGTGTGAGCAATATTGGAGTCTGATGATGATGATGATGATG 1809
DB 1441 ATGACTGAAATTTTGTGTGAGCAATATTGGAGTCTGATGATGATGATGATGATG 1500
QY 1810 AAGGCACTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1869
DB 1501 AAGGCACTGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
QY 1870 AAGAGGATGATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1929
DB 1561 AAGAGGATGATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1930 TTGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1989
DB 1621 TTGCGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1680
QY 1990 AATTTTATGAGTTTAAATTAAGAAATTTTACAA 2028
DB 1681 AATTTTATGAGTTTAAATTAAGAAATTTTACAA 1719

RESULT 9
US-60-453-050-274
: Sequence 274, Application US/60453050
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: LUKE, May
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001457
: CURRENT APPLICATION NUMBER: US/60/453,050
: NUMBER OF FILING DATE: 2003-03-10
: NUMBER OF SEQ ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 274
: LENGTH: 1719
: TYPE: DNA
: ORGANISM: Homo sapiens
US-60-453-050-274

Query Match 66.3%; Score 1353.8; DB 11; Length 1719;
Best Local Similarity 94.3%; Pred. No. 0;
Matches 1451; Conservative 6; Mismatches 3; Indels 79; Gaps 2;
QY 568 CAGAGCTTGCAGTTTGAAGAAACAGATCTCTCATAGTATGATTTCTCCATCGTCA 627

```

```

DB 182 CAGAGCTTGCAGTTTGAAGAAACAGATCTCTCATAGTATGATTTCTCCATCGTCA 241
QY 628 ACATACGAGAAAGATGCTGCTCCCTGAGAGTCTCTGAAATCTTAACCTTGCAT 687
DB 242 ACATACGAGAAAGATGCTGCTCCCTGAGAGTCTCTGAAATCTTAACCTTGCAT 301
QY 688 CATTGAGCTTACATGCTGACAGCACTTCTAGATGAGCCCTTGGAATAGAGGTGATGA 747
DB 302 CATTGAGCTTACATGCTGACAGCACTTCTAGATGAGCCCTTGGAATAGAGGTGATGA 361
QY 748 CATGCTGCTGCTGAGCTCTTGGAGAACTAAGAGATTTTCAAGAGCTATTAATAGA 807
DB 362 CATGCTGCTGCTGAGCTCTTGGAGAACTAAGAGATTTTCAAGAGCTATTAATAGA 421
QY 808 GAATGTTTGGAAATATGACAGAGATATCAATGATGATGATGATGATGATGATGATG 867
DB 422 GAATGTTTGGAAATATGACAGAGATATCAATGATGATGATGATGATGATGATGATG 481
QY 868 GGAATTTACAGAGTACCTGATCACTGAGGAAATGACAGAGAGTATGATGATGATG 927
DB 482 GGAATTTACAGAGTACCTGATCACTGAGGAAATGACAGAGAGTATGATGATGATG 541
QY 928 TCTAAGAAAG----- 938
DB 542 TCTAAGAAAGATATTGAGGCTGTGAACAGATTAAGATATGATGATGATGATGATG 601
QY 939 -----GACTGCTCCATTTGCTCCCTGCTGATG 969
DB 602 CCGATGCTTAAACACGTTTGTGAGCAGACCTGCTCCATTTGCTCCCTGCTGATG 661
QY 970 CTTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1029
DB 662 CTTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 721
QY 1030 CGTGAATTTTATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1089
DB 722 YGTGCAAAATTTTATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 781
QY 1090 TGAAGCTGAGAACAGAGAGATTAATCAGCTGTGTGATGATGATGATGATGATGATG 1149
DB 782 TGAAGCTGAGAACAGAGAGATTAATCAGCTGTGTGATGATGATGATGATGATGATG 841
QY 1150 GGAATGCGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1209
DB 842 GGAATGCGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 901
QY 1210 AAGAGATGAGACTGAGATGCTGTCAACACATAGAAATGATGATGATGATGATGATG 1269
DB 902 AAGAGATGAGACTGAGATGCTGTCAACACATAGAAATGATGATGATGATGATGATG 961
QY 1270 TAAACAAGAACTTAAAGGAGAGGATCCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1329
DB 962 TAAACAAGAACTTAAAGGAGAGGATCCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1021
QY 1330 CGTTAAGACATGTCAGTGGCCCTTTGAAAGTGCTGGGATCCAGTTCAGAGGA 1389
DB 1022 CGTTAAGACATGTCAGTGGCCCTTTGAAAGTGCTGGGATCCAGTTCAGAGGA 1081
QY 1390 GGGGATGTTTGTCTCCAGTCTGGGCAAGAGTCTATGCGGAATCTCTGCAATGA 1449
DB 1082 GGGGATGTTTGTCTCCAGTCTGGGCAAGAGTCTATGCGGAATCTCTGCAATGA 1141
QY 1450 GAAAGCCTCCACCATTTCTTTGATGTGCTTTAAAGTCTCAGTCTCTTAATAGA 1509
DB 1142 GAAAGCCTCCACCATTTCTTTGATGTGCTTTAAAGTCTCAGTCTCTTAATAGA 1201
QY 1510 AACAGAGCTCTGTGAGCTCTGTGCTGATGTCTGGAATGATGATGATGATGATGATG 1569
DB 1202 AACAGAGCTCTGTGAGCTCTGTGCTGATGTCTGGAATGATGATGATGATGATGATG 1261
QY 1570 AAAGCATTTTCTTTCTTTGAACTTAAAGTCTTATTAATTAAGACAGACAGATTC 1629

```



Db 722 CGTCAAGTTTCATCCAGGCGCCACAGAGCTACAGCCATCCATAAAGATGCCATACAA 781  
 Qy 1090 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATCGAATCATCATGG 1149  
 Db 782 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATCGAATCATCATGG 841  
 Qy 1150 GATCCCGAATGGGCGGCGACACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 1209  
 Db 842 GGATCCCGAATGGGCGGCGACACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 901  
 Qy 1210 AAGAAGATTGAGACTGAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTA 1269  
 Db 902 AAGAAGATTGAGACTGAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTA 961  
 Qy 1270 TAACAAAGAACTTAAGAGAGGATCCCGAGGCGAGAAATGATCAGAGCTGAATGACAG 1329  
 Db 962 TAACAAAGAACTTAAGAGAGGATCCCGAGGCGAGAAATGATCAGAGCTGAATGACAG 1021  
 Qy 1330 CGTTTAAGAGACATGTCACAGTGGCTTTGGAAAGTGGTGGGATCCAGTTCAGAGGGA 1389  
 Db 1022 CGTTTAAGAGACATGTCACAGTGGCTTTGGAAAGTGGTGGGATCCAGTTCAGAGGGA 1081  
 Qy 1390 GGGGTATGTTGTCTCCAGTCTGGGCAAGAGAGTGTATGCGGAATTCCTCTCATAGCA 1449  
 Db 1082 GGGGTATGTTGTCTCCAGTCTGGGCAAGAGAGTGTATGCGGAATTCCTCTCATAGCA 1141  
 Qy 1450 GAAAGCTCCACCATTTTCTTTGATGTTGATTAAGTCTACAGTCTCTTAATAGA 1509  
 Db 1142 GAAAGCTCCACCATTTTCTTTGATGTTGATTAAGTCTACAGTCTCTTAATAGA 1201  
 Qy 1510 AACAGAGTCTGTGACGCTGTGTGTGCTGATGTCTGGAATGATGATGTTGACG 1569  
 Db 1202 AACAGAGTCTGTGACGCTGTGTGTGCTGATGTCTGGAATGATGATGTTGACG 1261  
 Qy 1570 AAGGATTTTCTTTCTTTGATGTTGATTAAGTCTATTTAAAGCAGACAGATGTC 1629  
 Db 1262 AAGGATTTTCTTTCTTTGATGTTGATTAAGTCTATTTAAAGCAGACAGATGTC 1321  
 Qy 1630 ACATTTTATACATGAGATCTTTGTGTGATTAACAGAGATTGACTGATCCCTTTA 1689  
 Db 1322 ACATTTTATACATGAGATCTTTGTGTGATTAACAGAGATTGACTGATCCCTTT- 1380  
 Qy 1690 AAGAAGTTTATGTCCTGACCTGTGGCTAAATATATATATTTCCAGATGCTTTGTAG 1749  
 Db 1381 AAGAAGTTTATGTCCTGACCTGTGGCTAAATATATATATTTCCAGATGCTTTGTAG 1440  
 Qy 1750 ATAGAGATTTTGTGAGCAGCATATTGGAGTTCTATATTGAGTGAATGGCAGGAA 1809  
 Db 1441 ATAGAGATTTTGTGAGCAGCATATTGGAGTTCTATATTGAGTGAATGGCAGGAA 1500  
 Qy 1810 AGGGCATCTCATTTGATGATTAAGTGAACCAAACTAGTCTCGAATTTCTACAGAGA 1869  
 Db 1501 AGGGCATCTCATTTGATGATTAAGTGAACCAAACTAGTCTCGAATTTCTACAGAGA 1560  
 Qy 1870 AGAGAGATTCAGACTGAGAGAGCTGTGACATAGGACTTAAGACCAAGACTTTGAAT 1929  
 Db 1561 AGAGAGATTCAGACTGAGAGAGCTGTGACATAGGACTTAAGACCAAGACTTTGAAT 1620  
 Qy 1930 TTGCGAGCTCTCATGTGTGATTTATATACAGCTGCTCTTCTATGAGTTCAAACTC 1989  
 Db 1621 TTGCGAGCTCTCATGTGTGATTTATATACAGCTGCTCTTCTATGAGTTCAAACTC 1680  
 Qy 1990 ATATTTTATGAGTTTAATTAAGAAAAATTTACAA 2028  
 Db 1681 ATATTTTATGAGTTTAATTAAGAAAAATTTACAA 1719

RESULT 8  
 US-60-453-135-274  
 ; Sequence 274, Application US/60453135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele  
 ; APPLICANT: IAKOUBOVA, Olgia

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001456  
 ; CURRENT APPLICATION NUMBER: US/60/453,135  
 ; CURRENT FILING DATE: 2003-03-10  
 ; NUMBER OF SEQ ID NOS: 82762  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 274  
 ; LENGTH: 1719  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-60-453-135-274  
 Query Match 66.3%; Score 1353.8; DB 11; Length 1719;  
 Best Local Similarity 94.3%; Pred. No. 0;  
 Matches 1451; Conservative 6; Mismatches 3; Indels 79; Gaps 2;  
 Qy 568 CAGAGCTTGCAAGTTTTCAGAAACAGAAATCTCTCATAGTAATTTCTCATGCTCA 627  
 Db 182 CAGAGCTTGCAAGTTTTCAGAAACAGAAATCTCTCATAGTAATTTCTCATGCTCA 241  
 Qy 628 ACATAGGGAAGAGTGGTGGTCCCTTGAGAGTCTCTGAGTCTTAACCTTGAT 687  
 Db 242 ACATAGGGAAGAGTGGTGGTCCCTTGAGAGTCTCTGAGTCTTAACCTTGAT 301  
 Qy 688 CCTTTGCTTCATGCTGACAGGACAGTTCTAGATGAGCGCTGGATTAAGAGGTGATGA 747  
 Db 302 CCTTTGCTTCATGCTGACAGGACAGTTCTAGATGAGCGCTGGATTAAGAGGTGATGA 361  
 Qy 748 CATGCTTGCTGCTGGGCTCTTGAGAGAGTAAAGATTTTTCACAGACGCTAATACGAA 807  
 Db 362 CATGCTTGCTGCTGGGCTCTTGAGAGAGTAAAGATTTTTCACAGACGCTAATACGAA 421  
 Qy 808 GAATGTTTGGAAATATGCGAGAGCTATCAATGATATCTTCAATCAATGGCTTCAA 867  
 Db 422 GAATGTTTGGAAATATGCGAGAGCTATCAATGATATCTTCAATCAATGGCTTCAA 481  
 Qy 868 GGAATTTTCAGAGTACCTGATCTGACAGGAGAAATGACACAGTGAAGTGAACAGCT 927  
 Db 482 GGAATTTTCAGAGTACCTGATCTGACAGGAGAAATGACACAGTGAAGTGAACAGCT 541  
 Qy 928 TCTTAAGAAG----- 938  
 Db 542 TCTTAAGAAGGATTTGAGAGCTCTGAACAAGTAACTAAGATATGCGGAAACAAA 601  
 Qy 939 -----GACCTGGTCCCATTTGTCCTCCCTGCTATGG 969  
 Db 602 CCGATGGGTTAAAAACCGTTTTTTGAGCAGACCTGGTCCCATTTGTCCTCCCTGCTATGG 661  
 Qy 970 CTTAGAGATTCATGATGTCGAGAGTGGAGAGTCTGTTCTTGAACCTGCTTGAAT 1029  
 Db 662 CTTAGAGATTCATGATGTCGAGAGTGGAGAGTCTGTTCTTGAACCTGCTTGAAT 721  
 Qy 1030 CGTGCAAGTTTTCATCCAGGCGCAGAGCCTTACAGCCATGCAATTAAGATGCCATCAA 1089  
 Db 722 YGTGCAAGTTTTCATCCAGGCGCAGAGCCTTACAGCCATGCAATTAAGATGCCATCAA 781  
 Qy 1090 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATGGAATCATATTGG 1149  
 Db 782 TGAAGCTGAGAACAGAGAGATTATACCTGTGTGACCTCTGTGATGGAATCATATTGG 841  
 Qy 1150 GGATCCCGAATGGGCGGCGACACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 1209  
 Db 842 GGATCCCGAATGGGCGGCGACACTATAATCCCAATCCCACTTAACCACTGAAGAAAG 901  
 Qy 1210 AAGAAGTTGAGACTGAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTA 1269  
 Db 902 AAGAAGTTGAGACTGAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTA 961  
 Qy 1270 TAACAAAGAACTTAAGAGAGGATCCCGAGGCGAGAAATGATCAGAGCTGAATGACAG 1329  
 Db 962 TAACAAAGAACTTAAGAGAGGATCCCGAGGCGAGAAATGATCAGAGCTGAATGACAG 1021

Db 481 TCAATTGGCTTCAAGAAATTTACGAGTACGTATCATCTGAGGAAATATGCACACTGAG 540  
 QY 914 ACTAGTACCACTCTTCTTAAAGAAAGC----- 939  
 Db 541 ACTAGTACCACTCTTCTTAAAGAAAGTATGAGGCTCTGAAACAGTAACTAGAGATAT 600  
 QY 940 -----ACCTGGTCCCATTTGTC 955  
 Db 601 GCCCGGAAACAAACGAGTGGGTTAAAAACGTTTTTTTGAGCAGACCTGGTCCCATTTGTC 660  
 QY 956 CCCCCCTGTATGCTTAGAGTATCTGATGTCGAAAGTGGAGAGAGTCTGTTCTTGA 1015  
 Db 661 CCCCCGTGTACGCGCTTAGAGTATCTGATGTCGAAAGTGGAGAGTCTGTTCTTGA 720  
 QY 1016 CCTGCTCTTGAATTCGTGCAAAAGTTTCATCCAGGCGCACAAAGCTTACAGCCACTCCATA 1075  
 Db 721 CCTGCTCTTGAATTCGTGCAAAAGTTTCATCCAGGCGCACAAAGCTTACAGCCACTCCATA 780  
 QY 1076 AAGATGCATACATGATGAGTGAAGAACAGAGAGTATGACCTGCTGAGCTCTGAT 1135  
 Db 781 AAGATGCATACATGATGAGTGAAGAACAGAGAGTATGACCTGCTGAGCTCTGAT 840  
 QY 1136 CGAATCATCATTTGGGATCGCGAATGGCGACGACATAAATCCAAATCCCACTTGAAC 1195  
 Db 841 CGAATCATCATTTGGGATCGCGAATGGCGACGACATAAATCCAAATCCCACTTGAAC 900  
 QY 1196 CAACCTGAAGAAAGAAAGTATGAGTCTGATGCTCTCAACCAATGAAAGCAAGT 1255  
 Db 901 CAACCTGAAGAAAGAAAGTATGAGTCTGATGCTCTCAACCAATGAAAGCAAGT 960  
 QY 1256 GTTTCCTCCAGACTATACAAAGAAACCTTAAAGGAAAGGATCCCGAGGCGAGAAATGATCAA 1315  
 Db 961 GTTTCCTCCAGACTATACAAAGAAAGCTTAAAGGAAAGGATCCCGAGGCGAGAAATGATCAA 1020  
 QY 1316 GAGCTGAATGCAGCGTTTAAAGACATGTCCTCACTGCGCTTGAAGAGTGGGATC 1375  
 Db 1021 GAGCTGAATGCAGCGTTTAAAGACATGTCCTCACTGCGCTTGAAGAGTGGGATC 1080  
 QY 1376 CAGTTCAAGGAGGAGGAGTATGTTTGTCTCCCACTCTGGGCAAGAGTGTATGCGGAA 1435  
 Db 1081 CAGTTCAAGGAGGAGGAGTATGTTTGTCTCCCACTCTGGGCAAGAGTGTATGCGGAA 1140  
 QY 1436 TTTCTCTCATAGCAGAAAGCTCCACCATTTTCTTTGATGATGTTTAAAGCTCAGC 1495  
 Db 1141 TTTCTCTCATAGCAGAAAGCTCCACCATTTTCTTTGATGATGTTTAAAGCTCAGC 1200  
 QY 1496 TTTCTCTATAATAGAAACAGAGGCTTGTGACGCTCTGTGTGAGTATGTCGAAA 1555  
 Db 1201 TTTCTCTATAATAGAAACAGAGGCTTGTGACGCTCTGTGTGAGTATGTCGAAA 1260  
 QY 1556 TGATGTAGTTCAGGAAAGCAATTTTCTTTGAAACCTTAAAGTTCTATTTAA 1615  
 Db 1261 TGATGTAGTTCAGGAAAGCAATTTTCTTTGAAACCTTAAAGTTCTATTTAA 1320  
 QY 1616 GCAGCAGAGATTCACATTTTATACATGAGGATCTTCTGTGTGATGTCGAGATG 1675  
 Db 1321 GCAGCAGAGATTCACATTTTATACATGAGGATCTTCTGTGTGATGTCGAGATG 1380  
 QY 1676 ACTGCATCCCTTTAAAGAAAGTTTATGTCCTGACTCTGGCTAAATTAATCTAATTTCC 1735  
 Db 1381 ACTGCATCCCTTTTAAAGAAAGTTTATGTCCTGACTCTGGCTAAATTAATCTAATTTCC 1439  
 QY 1736 AGATGCTTTTGTAGATGATGAGTATTTGTGAGCCACATATTTGGAGTTCTAGATTGGA 1795  
 Db 1440 AGATGCTTTTGTAGATGATGAGTATTTGTGAGCCACATATTTGGAGTTCTAGATTGGA 1499  
 QY 1796 GTGATGGGCAAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAATTAATTTCCG 1855  
 Db 1500 GTGATGGGCAAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAATTAATTTCCG 1559  
 QY 1856 GAATTTACAGAGAGAGGAGGATCAGACTGAGAGAGCTGACATGAGACTTGAAGACC 1915  
 Db 1560 GAATTTACAGAGAGAGGAGGATCAGACTGAGAGAGCTGACATGAGACTTGAAGACC 1619

QY 1916 AAGACTTGAATTTGCGAGCTGCTCATGCTGAGTATATACAGTCTCTTCTAT 1975  
 Db 1620 AAGACTTGAATTTGCGAGCTGCTCATGCTGAGTATATACAGTCTCTTCTAT 1679  
 QY 1976 TGAGTTCAAAATCTATATTTTATTTGAAGTTTAAATGAAGAAATTTTCAAGAAAAA 2035  
 Db 1680 TGAGTTCAAAATCTATATTTTATTTGAAGTTTAAATGAAGAAATTTTCAAGAAAAA 1739  
 QY 2036 AAAAAA 2041  
 Db 1740 AAAAAA 1745  
 RESULT 7  
 US-10-170-235-41123  
 ; Sequence 41123, Application US/10170235  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig  
 ; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF  
 ; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREO  
 ; FILE REFERENCE: CL001380  
 ; CURRENT APPLICATION NUMBER: US/10/170, 235  
 ; CURRENT FILING DATE: 2003-03-17  
 ; NUMBER OF SEQ ID NOS: 42514  
 ; SEQ ID NO 41123  
 ; LENGTH: 1719  
 ; TYPE: DNA  
 ; ORGANISM: HUMAN  
 US-10-170-235-41123  
 Query Match 66.4%; Score 1354.6; DB 8; Length 1719;  
 Best Local Similarity 94.6%; Pred. No. 0;  
 Matches 1456; Conservative 0; Mismatches 4; Indels 79; Gaps 2;  
 QY 568 CAGAGGCTTGCAGATTTTGAAGAAACAGAAATCTCATAGTGAATTTCTCATGTCGA 627  
 Db 182 CAGAGGCTTGCAGATTTTGAAGAAACAGAAATCTCATAGTGAATTTCTCATGTCGA 241  
 QY 628 ACATACGAGAGAGAGTGTGTCCTCCCTTGGAGGTCCTCTGAAGTCTCTAACCCTTGAT 687  
 Db 242 ACATACGAGAGAGAGTGTGTCCTCCCTTGGAGGTCCTCTGAAGTCTCTAACCCTTGAT 301  
 QY 688 CCTTGGCTCATGCTGACCGAGGAGTCTATGATGAGCGCTTGATGAAGAGGAGTGA 747  
 Db 302 CCTTGGCTCATGCTGACCGAGGAGTCTATGATGAGCGCTTGATGAAGAGGAGTGA 361  
 QY 748 CATGCTTGTGCTGGGCTCTTGGAGAACTPAAGATTTTCAAGAGCTTAATTAATCAGAA 807  
 Db 362 CATGCTTGTGCTGGGCTCTTGGAGAACTPAAGATTTTCAAGAGCTTAATTAATCAGAA 421  
 QY 808 GAATGTTTGGAAATTAAGCAGAGACTATCAACATGATCTTCCATCAATTTGGCTTCAA 867  
 Db 422 GAATGTTTGGAAATTAAGCAGAGACTATCAACATGATCTTCCATCAATTTGGCTTCAA 481  
 QY 868 GGAATTTACAGAGTACCTGATCACTGAGGAGAAATGACACTGAGAGCTAGTAACGAGCT 927  
 Db 482 GGAATTTACAGAGTACCTGATCACTGAGGAGAAATGACACTGAGAGCTAGTAACGAGCT 541  
 QY 928 TCTAAGAAAG----- 938  
 Db 542 TCTAAGAAAGTATGAGGCTCTGAACAAGTAAAGATATGCCGGAAACAAA 601  
 QY 939 -----GACCTGTGCCATTTGTCCTCCCTGCTATG 969  
 Db 602 CCGATGGGTTAAAAACGTTTTTTGAGCAGACCTGTGCCATTTGCCCTGCTATG 661  
 QY 970 CTTAGAGTATCTATGCTCTGAAGTGGGAGAGTCTGTCTTAACCTGCTTGAAT 1029  
 Db 662 CTTAGAGTATCTATGCTCTGAAGTGGGAGAGTCTGTCTTAACCTGCTTGAAT 721  
 QY 1030 CGTCAAGAGTTTCTCAAGGCGCACAGCCATCAAGCCACATCAATTAAGATGCCATCAA 1089



QY	1229	GCATCAACACCCATTAAGAGTCAGAGTGTTCCTCCAGACCTATTAACAAAGAACCTPAAGG	1289
Db	1321	GCCTGACACACCTAGTAAGAGTCAGAGTGTTCCTCCAGACCATTAACAAAGAACCTPAAGG	1380
QY	1289	AAGGAGATCCCCAGAGCGAGATGATCAAGAGCTGAAGATGACGCGTTTAAGAGACATGTCCA	1348
Db	1381	AAGGAGATCCCCAGAGCGAGATGATCAAGAGCTGAAGATGACGCGTTTAAGAGACATGTCCA	1440
QY	1349	GTCGCTTTTGAAGAGTGTGGGGATCCAGTTCAGAGAGGAGGGGTATGTTTGTCTCCA	1408
Db	1441	GTCGCTTTTGAAGAGTGTGGGGATCCAGTTCAGAGAGGAGGGGTATGTTTGTCTCCA	1500
QY	1409	GTCCTGGGCAAGAGATGCTATCGCGAATCTCTGCATATGACGAAAAAGCTCCACCATTTT	1468
Db	1501	GTCCTGGGCAAGAGATGCTATCGCGAATCTCTGCATATGACGAAAAAGCTCCACCATTTT	1560
QY	1469	CTTTTGTGAGGTTTTAAAGTCTCAGCTCTCTATTAATTAAGAAACGACAGCTTGTACGC	1528
Db	1561	CTTTTGTGAGGTTTTAAAGTCTCAGCTCTCTATTAATTAAGAAACGACAGCTTGTACGC	1620
QY	1529	TCCTTGTGTGCGATGTGTCTGGAAATGATGTAGTTCAGAAACATTTTTTTTTTCTT	1588
Db	1621	TCCTTGTGTGCGATGTGTCTGGAAATGATGTAGTTCAGAAACATTTTTTTTTTCTT	1680
QY	1589	TGAACCTTAAGGTTCTATTAATTAAGAACGACAGATTCACATTTTATATCATAGAGA	1648
Db	1681	TGAACCTTAAGGTTCTATTAATTAAGAACGACAGATTCACATTTTATATCATAGAGA	1740
QY	1649	TCTTCTTTGGGGAATACAGAGATGACCTGCATCCCTTAAGAAGATTATATGTCCCT	1708
Db	1741	TCTTCTTTGGGGAATACAGAGATGACCTGCATCCCTTAAGAAGATTATATGTCCCT	1799
QY	1709	GACTCTGGCTAAATTAATTAATTTTCAGATGCTTTTGTAGATGACTGAAGTATTGTGA	1768
Db	1800	GACTCTGGCTAAATTAATTAATTTTCAGATGCTTTTGTAGATGACTGAAGTATTGTGA	1859
QY	1769	GCCACATATTGGGAGTTCTAGATTTAGTGATGCGCAGAAAGGCCATCTCATTTGAGA	1828
Db	1860	GCCACATATTGGGAGTTCTAGATTTAGTGATGCGCAGAAAGGCCATCTCATTTGAGA	1919
QY	1829	TGATTAAGTAAACCAACATGTTCTCGGAATCTACAGAGAAGGAGGATCAGACTAG	1888
Db	1920	TGATTAAGTAAACCAACATGTTCTCGGAATCTACAGAGAAGGAGGATCAGACTAG	1979
QY	1889	GAAGCTGTGACATAGGACTTGAAGACCAAGATTGAATTTGGCAGCTGCTCATGTGT	1948
Db	1980	GAAGCTGTGACATAGGACTTGAAGACCAAGATTGAATTTGGCAGCTGCTCATGTGT	2033
QY	1949	GAGTTATTAATCAGCTGCTCTTCTATATGAGTTACAATCTATATTTTATGAAGTTTA	2008
Db	2040	GAGTTATTAATCAGCTGCTCTTCTATATGAGTTACAATCTATATTTTATGAAGTTTA	2099
QY	2009	AATTAAGAAAAAATTTACAA 2028	
Db	2100	AATTAAGAAAAAATTTACAA 2119	
RESULT 5			
US-60-453-050-275			
Sequence 275, Application US/60453050			
GENERAL INFORMATION:			
APPLICANT: CARGILL, Michele			
APPLICANT: LUKE, May			
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
TITLE OF INVENTION: STENOISIS, METHODS OF DETECTION AND USES THEREOF			
FILE REFERENCE: CL001457			
CURRENT APPLICATION NUMBER: US/60/453,050			
NUMBER OF SEQ ID NOS: 82762			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 275			
LENGTH: 2119			
TYPE: DNA			

ORGANISM: Homo sapiens  
US-60-453-050-275

Query Match	91.6%;	Score 1869.6;	DB 11;	Length 2119;
Best Local Similarity	94.4%;	Pred. No. 0;		
Matches 2002; Conservative	8;	Mismatches	7;	Indels 103; Gaps. 3;

OY	11	ATGCGTCGCCGGGGGGCTGCACAGAGACTTCCTGTGGGACAGTGGGCTCAGAGGGGCTCCAA	70
Db	1	ATGGCGTCCGGGGGGCTGCACAGAGACTTCCTGTGGGACAGTGGGCTCAGAGGGGCTCCAA	60
OY	71	CGGACCCCTACCTCTTGTAGTGATTTCTCGGGGGCCACGGGCAACGGCAAAATCCAGCTGGCG	130
Db	61	CGGACCCCTACCTCTTGTAGTGATTTCTCGGGGGGGCCACGGGCAACGGCAAAATCCAGCTGGCG	120
OY	131	TTGCAGCTAGGCCACGGGGCTCGGGGGTGAGATCGTCAAGCGCTGACTCTCATGCAAGTGTAT	190
Db	121	TTGCAGCTAGGCCACGGGGCTCGGGGGTGAGATCGTCAAGCGCTGACTCTCATGCAAGTGTAT	180
OY	191	GAAGGCGCTAGACATCATCACCAACAAGGTTTCTGCCAAGAGCAGAGAATCTGGCGGCAC	250
Db	181	GAAGGCGCTAGACATCATCACCAACAAGGTTTCTGCCAAGAGCAGAGAATCTGGCGGCAC	240
OY	251	CACATGATCAGCTTTGGTGATCTCTTGTGACCAATTACACAGTGTGGACTTCAGAAAT	310
Db	241	CACATGATCAGCTTTGGTGATCTCTTGTGACCAATTACACAGTGTGGACTTCAGAAAT	300
OY	311	AGAGCACTGCTCTGATTGAA-----GATATATTGGCCCGA	346
Db	301	AGAGCACTGCTCTGATTGAA-----GATATATTGGCCCGA	366
OY	347	GACAAATTCCTATTGGTTGTGGAGAGAAACAAATTATTACATTGAATCTGCTGTGGAAA	406
Db	361	GACAAATTCCTATTGGTTGTGGAGAGAAACAAATTATTACATTGAATCTGCTGTGGAAA	422
OY	407	GTTCTTGTCAATACCAAGCCCCAGAGAGATGGGGCACTGAGAAAGTGAATTGACCGAGAAATG	466
Db	421	GTTCTTGTCAATACCAAGCCCCAGAGAGATGGGGCACTGAGAAAGTGAATTGACCGAGAAATG	486
OY	467	GAGCTTGAAAAGAGAGATGGGCTTGTACTTACCAAGCGCTTAACCGAGTGGACCCAGAA	522
Db	481	GAGCTTGAAAAGAGAGATGGGCTTGTACTTACCAAGCGCTTAACCGAGTGGACCCAGAA	540
OY	527	ATGGCTGCCAAGCTGCATCCACATGACAAAGCAAAAGCGCCAGAGACTTCCAAGTTT	586
Db	541	ATGGCTGCCAAGCTGCATCCACATGACAAAGCAAAAGCGCCAGAGACTTCCAAGTTT	600
OY	587	GAAGAAACAGGAATCTCTCATAGTGAATTTCTCATGCTCAACTACGAGGAAGAGTGT	646
Db	601	GAAGAAACAGGAATCTCTCATAGTGAATTTCTCATGCTCAACTACGAGGAAGAGTGT	666
OY	647	GGTCCCTTGGAGGTCTCTCTAAGTTCTCTAACCCTTGATCCTTTGGCTTCATGCTGAC	706
Db	661	GGTCCCTTGGAGGTCTCTCTAAGTTCTCTAACCCTTGATCCTTTGGCTTCATGCTGAC	720
OY	707	CAGGAGTTTATGATGAGACGGCTTGGATTAGAAGGGTGATGACATGCTGCTGGGCTC	766
Db	721	CAGGAGTTTATGATGAGACGGCTTGGATTAGAAGGGTGATGACATGCTGCTGGGCTC	786
OY	767	TTGAGGAAGTAAGAGATTTTACAGACGCTATAATCAGAAAGTGTTCGGAAATAGC	822
Db	781	TTGAGGAAGTAAGAGATTTTACAGACGCTATAATCAGAAAGTGTTCGGAAATAGC	844
OY	827	CAGGACTTCAACATGATGATCTTCCAATCATTTGGCTTCAAGGAATTTCAAGACTACTCG	888
Db	841	CAGGACTTCAACATGATGATCTTCCAATCATTTGGCTTCAAGGAATTTCAAGACTACTCG	900
OY	887	ATCACTGAGGAAAAATGCACACTGGAGACTAGTAACCAAGCTTCTTAAAGAAAG-----	936
Db	901	ATCACTGAGGAAAAATGCACACTGGAGACTAGTAACCAAGCTTCTTAAAGAAAGATTTAG	966
OY	939	-----	936

```
Db 1621 TCCTTTGTGCTGATGTCTGGAATGATGATTCAGGAAAGCATTTTCTT 1680
Oy 1589 TGAACCTTAAGCTTCTATATTAATAAGCAGCAGATTCACATTTTATACATGAGA 1648
Db 1681 TGAACCTTAAGCTTCTATATTAATAAGCAGCAGATTCACATTTTATACATGAGA 1740
Oy 1649 TCTTCTTTGTGGAATACAGATGATGATGATTCCTTTTAAAGAGTTTATGCTCT 1708
Db 1741 TCTTCTTTGTGGAATACAGATGATGATGATTCCTTTTAAAGAGTTTATGCTCT 1799
Oy 1709 GACTGTGCTTAATAATATCTAATTCACATGCTTTTGTGATGATGATGATGATGATGAT 1768
Db 1800 GACTGTGCTTAATAATATCTAATTCACATGCTTTTGTGATGATGATGATGATGATGAT 1859
Oy 1769 GCCACATATTGGAGTTCTAGATTTGAGTAAATGGCAGGAAGGCCATCTCCATTGAGA 1828
Db 1860 GCCACATATTGGAGTTCTAGATTTGAGTAAATGGCAGGAAGGCCATCTCCATTGAGA 1919
Oy 1829 TGATTAAGTAAACCAACTGTTCTGGATTTCTACAGAAAGAGGATCAGACTGAG 1888
Db 1920 TGATTAAGTAAACCAACTGTTCTGGATTTCTACAGAAAGAGGATCAGACTGAG 1979
Oy 1889 GAAGCTGTGACATAGGACTGGAAGACCAAGACTTTGAATTTGCGAGCTGCTCATGTGT 1948
Db 1980 GAAGCTGTGACATAGGACTGGAAGACCAAGACTTTGAATTTGCGAGCTGCTCATGTGT 2039
Oy 1949 GACTTATATACAGTGTCTTCTATATGATTAACAATCTAATTTTATGAGTTTA 2008
Db 2040 GACTTATATACAGTGTCTTCTATATGATTAACAATCTAATTTTATGAGTTTA 2099
Oy 2009 AATAAGAAAAATTTACAA 2028
Db 2100 AATAAGAAAAATTTACAA 2119

RESULT 4
US-60-453-135-275
; Sequence 275, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001436
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 275
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-275

Query Match 91.6%; Score 1869.6; DB 11; Length 2119;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 2002; Conservative 8; Mismatches 7; Indels 103; Gaps 3;
```

```
Oy 251 CACATGATCAGCTTTGTGATTCCTCTTGATACCAATTAACAGTGTGGACTTACGAAT 310
Db 241 CACATGATCAGCTTTGTGATTCCTCTTGATACCAATTAACAGTGTGGACTTACGAAT 300
Oy 311 AGAGCAACTGCTCGATTGAA-----GATATATTGGCCGA 346
Db 301 AGAGCAACTGCTCGATTGAA-----GATATATTGGCCGA 360
Oy 347 GACAAATTCCTATTGTTGGAGAGAACCAATTAATTAATTAATTAATTAATTAATTAATTA 406
Db 361 GACAAATTCCTATTGTTGGAGAGAACCAATTAATTAATTAATTAATTAATTAATTAATTA 420
Oy 407 GTTCTTGTCAATTAACCAAGCCCGAGAGATGGCCTGAGAAAGTATGATGACCCGAAAGTG 466
Db 421 GTTCTTGTCAATTAACCAAGCCCGAGAGATGGCCTGAGAAAGTATGATGACCCGAAAGTG 480
Oy 467 GAGCTGAAAGAGAGATGGTCTGTGATCTTCAACCAAGCCCTAAGCCAGGTGGACCGAA 526
Db 481 GAGCTGAAAGAGAGATGGTCTGTGATCTTCAACCAAGCCCTAAGCCAGGTGGACCGAA 540
Oy 527 ATGGCTCCCAAGCTGCATTCACATGACAAACGCAAAAGTGCCAGAGCTTGCAAGTTT 586
Db 541 ATGGCTCCCAAGCTGCATTCACATGACAAACGCAAAAGTGCCAGAGCTTGCAAGTTT 600
Oy 587 GAAGAAACAGGAATCTCATAGTGAATTTTCCATGCTCAGATCAGGAAGAGTGT 646
Db 601 GAAGAAACAGGAATCTCATAGTGAATTTTCCATGCTCAGATCAGGAAGAGTGT 660
Oy 647 GGTCCCTTGAGAGTCTCTGAAATCTCTAACCCTTGCACTTGGCTTCAAGCTGAC 706
Db 661 GGTCCCTTGAGAGTCTCTGAAATCTCTAACCCTTGCACTTGGCTTCAAGCTGAC 720
Oy 707 CAGGAGTTCATAGATGAGCGCTTGATTAAGAGGTGATGATGATGATGATGATGATGATG 766
Db 721 CAGGAGTTCATAGATGAGCGCTTGATTAAGAGGTGATGATGATGATGATGATGATGATG 780
Oy 767 TTGAGGAACTAAGAGATTTTACAGACGCTTAATTAACGAAGATTTTCCGAAATAGC 826
Db 781 TTGAGGAACTAAGAGATTTTACAGACGCTTAATTAACGAAGATTTTCCGAAATAGC 840
Oy 827 CAGGACTATCAACATGGATCTTCCAAATCAATTTGGCTTCAAGGAATTTACAGAGTACTG 886
Db 841 CAGGACTATCAACATGGATCTTCCAAATCAATTTGGCTTCAAGGAATTTACAGAGTACTG 900
Oy 887 ATCACTGAGGAAATGACACTGAGAGCTAGTAACCAAGCTTCTTAAGAAAG----- 938
Db 901 ATCACTGAGGAAATGACACTGAGAGCTAGTAACCAAGCTTCTTAAGAAAGATTTGAG 960
Oy 939 ----- 938
Db 961 GCTCTGAACCAAGTAACATAAGATATGCCCGAAACAAACCGATGGTTAAAAACCGT 1020
Oy 939 -----GACCTGGCCATATGTCCTCCCTGCTATGAGTTAGATATGATGATGATGATG 988
Db 1021 TTTTGAACAACCTGTGCTCCATTTGTCCTCCCTGCTATGATGATGATGATGATGATGATG 1080
Oy 989 TCGAAGTGGAGAGTCTGTTTGAACCTCTCTTGAATCTGCAAGTCTTCAAGTTCACAG 1048
Db 1081 TCGAAGTGGAGAGTCTGTTTGAACCTCTCTTGAATCTGCAAGTCTTCAAGTTCACAG 1140
Oy 1049 GGCCACAAGCCTTACAGCCTCAATAAAGATGCCATCAATGAAGCTGAGAACAGAGA 1108
Db 1141 GGCCACAAGCCTTACAGCCTCAATAAAGATGCCATCAATGAAGCTGAGAACAGAGA 1200
Oy 1109 AGTTATCACTGTGTGACCTTGTGATGAATCATATTTGGGATCGGGAATGGCAGCG 1168
Db 1201 AGTTATCACTGTGTGACCTTGTGATGAATCATATTTGGGATCGGGAATGGCAGCG 1260
Oy 1169 CACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1228
Db 1261 CACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1320
```

```

Db 1939 TTCGAAATTTCTACAGAGAGAGGGAATCAGACTGAGAGAGCTGTGACATAGACTTG 1998
QY 1910 AAGACCAAGAGCTTTGAAATTTTGCAGAGCTGCTCATGTTGATATATATACAGCTGCT 1969
    |||||
Db 1999 AAGACCAAGAGCTTTGAAATTTTGCAGAGCTGCTCATGTTGATATATATACAGCTGCT 2058
QY 1970 TTCTATTGAGTTACAAATCTATATTTTATTTATTTGAAGTTTAAATAAGAAAAATTTTACAG 2029
    |||||
Db 2059 TTCTATTGAGTTACAAATCTATATTTTATTTATTTGAAGTTTAAATAAGAAAAATTTTACAG 2118
QY 2030 AAAAAAAAAA 2041
    |||||
Db 2119 AAAAAAAAAA 2130

RESULT 3
US-10-170-235-40554
; Sequence 40554, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 40554
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-40554

Query Match          91.6%; Score 1869.6; DB 8; Length 2119;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 2008; Conservative 0; Mismatches 9; Indels 103; Gaps 3;

QY 11 AAGGCTCCGCTGAGGCTGACAGAGAGAGTCTGTGGGAGAGGGGCTCAGGGGCTGCA 70
    |||||
Db 1 AAGGCTCCGCTGAGGCTGACAGAGAGAGTCTGTGGGAGAGGGGCTCAGGGGCTGCA 60
QY 71 CGGACCTACCTCTTGTAGTATTCGCGGCGCAGCGGCAACGCAATCCAGCGTGGC 130
    |||||
Db 61 CGGACCTACCTCTTGTAGTATTCGCGGCGCAGCGGCAACGCAATCCAGCGTGGC 120
QY 131 TTGCACTAGGCGGAGGCTCGGCGGTGAGATGTCAAGCGCTCAGTCCATGCAAGTCTAT 190
    |||||
Db 121 TTGCACTAGGCGGAGGCTCGGCGGTGAGATGTCAAGCGCTCAGTCCATGCAAGTCTAT 180
QY 191 GAAGGCTAGACATCATCAACAAGGTTTCTGCCAAGAGAGAGAAATCTGCGGCGAC 250
    |||||
Db 181 GAAGGCTAGACATCATCAACAAGGTTTCTGCCAAGAGAGAGAAATCTGCGGCGAC 240
QY 251 CACATGATCAGCTTTGTGATCTCTTGTGACCAATTTACACAGTGTGAGCTTCAGAAAT 310
    |||||
Db 241 CACATGATCAGCTTTGTGATCTCTTGTGACCAATTTACACAGTGTGAGCTTCAGAAAT 300
QY 311 AGAGCAACTGCTCTGATTTGAA-----GATATATTTGCCCA 346
    |||||
Db 301 AGAGCAACTGCTCTGATTTGAA-----GATATATTTGCCCA 360
QY 347 GACAAATTTCTATTGTTGTGGAGAGACCAATTTATTTACATTTGATCTGCTGAGAA 406
    |||||
Db 361 GACAAATTTCTATTGTTGTGGAGAGACCAATTTATTTACATTTGATCTGCTGAGAA 420
QY 407 GTTCTTTGTAATACCAAGCCCAAGAGATGGGCACTGAGAAATGATTTGACGGAAGTG 466
    |||||
Db 421 GTTCTTTGTAATACCAAGCCCAAGAGATGGGCACTGAGAAATGATTTGACGGAAGTG 480
QY 467 GAGCTTGAAGAGAGATGTTGTGATTTCAAAACGCTTAGCCAGGAGGAGCCAGAA 526
    |||||
Db 481 GAGCTTGAAGAGAGATGTTGTGATTTCAAAACGCTTAGCCAGGAGGAGGAGCCAGAA 540
QY 527 ATGGCTGCCAAGCTGCATCCATGACAAACGCAAGGTGGCCAGAGCTTGAAGTTT 586

```

```

Db 541 ATGGCTGCCAAGCTGCATCCATGACAAACGCAAAATGGCCAGAGCTTGAAGTTT 600
QY 587 GAAGAAACAGAAATCTCTCATAGTGAATTTCTCATGCTGATACATACGGAAGAGTGCT 646
    |||||
Db 601 GAAGAAACAGAAATCTCTCATAGTGAATTTCTCATGCTGATACATACGGAAGAGTGCT 660
QY 647 GGTCCCTTGGAGGCTCTGGAAGTTCTTAACCTTGCATCCCTTGGCTTCATGCTGAC 706
    |||||
Db 661 GGTCCCTTGGAGGCTCTGGAAGTTCTTAACCTTGCATCCCTTGGCTTCATGCTGAC 720
QY 707 CAGGCACTTCTAGATGAGCCCTTGATTAAGAGGAGGATGACATGCTTCTGCTGGGCTC 766
    |||||
Db 721 CAGGCACTTCTAGATGAGCCCTTGATTAAGAGGAGGATGACATGCTTCTGCTGGGCTC 780
QY 767 TTGGAGGAGACTAAGAGATTTTTCACAGAGCTATTAATCAGAAAGATGTTTCGAAATAGC 826
    |||||
Db 781 TTGGAGGAGACTAAGAGATTTTTCACAGAGCTATTAATCAGAAAGATGTTTCGAAATAGC 840
QY 827 CAGGCACTAAGAGATGATCTCTCCATCAATGCGTTCAAGGAATTTCAAGAGTACCTG 886
    |||||
Db 841 CAGGCACTAAGAGATGATCTCTCCATCAATGCGTTCAAGGAATTTCAAGAGTACCTG 900
QY 887 ATCACTAGGAGAAATGACACTGAGAGACTAGTAAACAGCTTCTTAAAGAAAG----- 938
    |||||
Db 901 ATCACTAGGAGAAATGACACTGAGAGACTAGTAAACAGCTTCTTAAAGAAAGTATTGAG 960
QY 939 ----- 938
Db 961 GCTCTGAACAGTAACATAAGATATGCCGGAACAAACCGATGGTTTAAACCGT 1020
QY 939 -----GACCTGCTCCATTTGTCGCCCTGTCTATGCGTTAGAGGTATGATGTC 988
    |||||
Db 1021 TTTTGAAGAGAGCTGCTCCATTTGCCCTGTCTATGCGTTAGAGGTATGATGTC 1080
QY 989 TCGAAGTGGAGAGAGTGTCTTTGAACCTGCTTGAATTCGCAAAATTTTCATCCAG 1048
    |||||
Db 1081 TCGAAGTGGAGAGAGTGTCTTTGAACCTGCTTGAATTCGCAAAATTTTCATCCAG 1140
QY 1049 GGCCACAAGCTACAGGCACTCCAAATTAAGATCCATTAAGAGTGAAGCAAGAGA 1108
    |||||
Db 1141 GGCCACAAGCTACAGGCACTCCAAATTAAGATCCATTAAGAGTGAAGCAAGAGA 1200
QY 1109 AGTTATCAGCTGTGTGACCTCTGTGATCAATATCATCTTTGGGATCGCAATGGGCAAGC 1168
    |||||
Db 1201 AGTTATCAGCTGTGTGACCTCTGTGATCAATATCATCTTTGGGATCGGCAATGGGCAAGC 1260
QY 1169 CACATAAATCCAAATCCCACTTGAACCAACTGAAGAAAGAAAGATGGACTAGAT 1228
    |||||
Db 1261 CACATAAATCCAAATCCCACTTGAACCAACTGAAGAAAGAAAGATGGACTAGAT 1320
QY 1229 GCTGTCAACACATAGAAAGTCAAGAGTGTTCGCCAGACTATTAACAAAGAACTTAAAGG 1288
    |||||
Db 1321 GCTGTCAACACATAGAAAGTCAAGAGTGTTCGCCAGACTATTAACAAAGAACTTAAAGG 1380
QY 1289 AAGGAGTCCCGCAGGAGAGATGATCAAGAGCTTAATGCAAGCTTTAAGAGATATGCCA 1348
    |||||
Db 1381 AAGGAGTCCCGCAGGAGAGATGATCAAGAGCTTAATGCAAGCTTTAAGAGATATGCCA 1440
QY 1349 GTGGCCTTTGGAAGAGTGTGGGAGATCCAGTTAGAGAGGAGGATATGTTTCTCCCA 1408
    |||||
Db 1441 GTGGCCTTTGGAAGAGTGTGGGAGATCCAGTTAGAGAGGAGGATATGTTTCTCCCA 1500
QY 1409 GTTGGGCAAGAGATGCTATGCGGAATTTCTGCAATAGCAGAAAGTCTCCACATTTT 1468
    |||||
Db 1501 GTTGGGCAAGAGATGCTATGCGGAATTTCTGCAATAGCAGAAAGTCTCCACATTTT 1560
QY 1469 CTTTATGTTGTTTAAAGTCTCAGCTTCTATATTAAGAAACACAGAGTCTGTGAC 1528
    |||||
Db 1561 CTTTATGTTGTTTAAAGTCTCAGCTTCTATATTAAGAAACACAGAGTCTGTGAC 1620
QY 1529 TCTTGTGTGCTGATGTCTGTGAATGATGATGTTCAAGGAAGCATTTTCTTCTT 1588
    |||||

```

PRIOR FILING DATE: 2000-09-11  
NUMBER OF SEQ ID NOS: 888  
SOFTWARE: Custom  
SEQ ID NO: 119  
LENGTH: 2130  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (191)..(1426)  
US-10-380-731-119

Query Match 94.1%; Score 1921.2; DB 8; Length 2130;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 2025; Conservative 0; Mismatches 8; Indels 79; Gaps 2;

QY 8 AAGATGGCGTCCGTGGCGGTGACAGACAGTTCCTGTGGGAGTGGGCTCAGAGGGCTTG 67  
DB 20 ACGATTCCGTCCGTGGCGGTGACAGACAGTTCCTGTGGGAGTGGGCTCAGAGGGCTTG 79  
QY 68 CAACGAGACCTACCTCTTGTAGTATCTCGGGCCACGGGACCGGCAATCCAGCTG 127  
DB 80 CAACGAGACCTACCTCTTGTAGTATCTCGGGCCACGGGACCGGCAATCCAGCTG 139  
QY 128 GCGTTCAGCTAGGCGAGCGGCTCGCGGTGAGATCGTCAAGCGCTGACTCCATGCAAGTTC 187  
DB 140 GCGTTCAGCTAGGCGAGCGGCTCGCGGTGAGATCGTCAAGCGCTGACTCCATGCAAGTTC 199  
QY 188 TATGAGGCGCTACATCATCATCAGCAACAGGCTTCGCGCCAGAGAGCAATCTGCCG 247  
DB 200 TATGAGGCGCTACATCATCATCAGCAACAGGCTTCGCGCCAGAGAGCAATCTGCCG 259  
QY 248 CACCATGATGACGCTTGTGTGATCTCTGTGACCAATTCACAGTGTGAGCTTCAGA 307  
DB 260 CACCATGATGACGCTTGTGTGATCTCTGTGACCAATTCACAGTGTGAGCTTCAGA 319  
QY 308 AATGAGCAACTGCTCTGATGTAAGATATATTTGCCGAGACAAATTCCTATTTGTTG 367  
DB 320 AATGAGCAACTGCTCTGATGTAAGATATATTTGCCGAGACAAATTCCTATTTGTTG 379  
QY 368 GAGGAGAACCAATTATATCATGATGATCTGCTCGGAAAGTTCTGTCATACCAACCCC 427  
DB 380 GAGGAGAACCAATTATATCATGATGATCTCTCTGGAATTTCTTGTCATACCAACCCC 439  
QY 428 CAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAGCTTGAAGAGAGATGAT 487  
DB 440 CAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAGCTTGAAGAGAGATGAT 499  
QY 488 CTGTACTTACAAAGCCCTTAAGCCAGAGTGACCCAGAAATGGCTGCCAAGCTGCATCCA 547  
DB 500 CTGTACTTACAAAGCCCTTAAGCCAGAGTGACCCAGAAATGGCTGCCAAGCTGCATCCA 559  
QY 548 CATGACAAAGCAAGTGGCGAGAGCTTGAAGTTTGAAGAAAGAGATCTCAT 607  
DB 560 CATGACAAAGCAAGTGGCGAGAGCTTGAAGTTTGAAGAAAGAGATCTCAT 619  
QY 608 AGTGAATTTCTCATCGTCAACATACGGAAGAGTGTGCTCCCTTGGAGTCTCTG 667  
DB 620 AGTGAATTTCTCATCGTCAACATACGGAAGAGTGTGCTCCCTTGGAGTCTCTG 679  
QY 668 AAGTTCCTAACCCTTGCATCTTGTGCTTCATGCTGACAGGAGTTCATGAGAGACGC 727  
DB 680 AAGTTCCTAACCCTTGCATCTTGTGCTTCATGCTGACAGGAGTTCATGAGAGACGC 739  
QY 728 TTGGATAGAGGGGATGATGATGCTGTGCTGGGCTCTTGGAGAACTAAGAGATTTT 787  
DB 740 TTGGATAGAGGGGATGATGATGCTGTGCTGGGCTCTTGGAGAACTAAGAGATTTT 799  
QY 788 CACAGAGCTATATATCAGAGAAATGTTTCGAAAAATAGCCAGACTATCAACATGATATC 847  
DB 800 CACAGAGCTATATATCAGAGAAATGTTTCGAAAAATAGCCAGACTATCAACATGATATC 859  
QY 848 TTCAATCAATTTGGCTTCAAGGAATTTACAGAGTACTGATGAGGCAAAATGCACA 907

DB 860 TTCCAAATCAATTTGGCTTCAAGGAATTTACAGATACCTGATCTAGGGAAAAATGCACA 919  
QY 908 CTGAGACTAGTATACAGAGCTTCTAAGAAAG----- 938  
DB 920 CTGAGACTAGTATACAGAGCTTCTAAGAAAGTATGAGGCTCTGAACAAGTAACTAAG 979  
QY 939 -----GACCTGTCC 949  
DB 980 AGATATGCCCGGAACAAACCGATGGTTAAACCGTTTTTGGAGACAGCTGTGCC 1039  
QY 950 ATTGTCCCCCTCTCTATGAGAGTATCTGATGTCTCGAAGGGGAGGAGTCTGT 1009  
DB 1040 ATTGTCCCCCTCTCTATGAGAGTATCTGATGTCTCGAAGGGGAGGAGTCTGT 1099  
QY 1010 CTGGAACCTGCTTGAATCGTCAAGTATTCATCCAGGGCCCAAGCTTACAGCACT 1069  
DB 1100 CTGGAACCTGCTTGAATCGTCAAGTATTCATCCAGGGCCCAAGCTTACAGCACT 1159  
QY 1070 CCAATAAAGATGCCATATCAATGAGAGCTGAGACAGAGAAAGTTATCCTGTGACTC 1129  
DB 1160 CCAATAAAGATGCCATATCAATGAGAGCTGAGACAGAGAAAGTTATCCTGTGACTC 1219  
QY 1130 TGTATCGATCATCATTTGGGATTCGGGATGGGACAGGCAATTAATCCAAATCCAC 1189  
DB 1220 TGTATCGATCATCATTTGGGATTCGGGATGGGACAGGCAATTAATCCAAATCCAC 1279  
QY 1190 TTGAACCACTGAAGAAAGAGAGATTTGAGCTGAGATGCTGCAACACCATAGAAAT 1249  
DB 1280 TTGAACCACTGAAGAAAGAGAGATTTGAGCTGAGATGCTGCAACACCATAGAAAT 1339  
QY 1250 CAGAGTGTTCCTCCAGCTATTAACAAGAACTTAAGGAAAGGATCCCAAGGCAAT 1309  
DB 1340 CAGAGTGTTCCTCCAGCTATTAACAAGAACTTAAGGAAAGGATCCCAAGGCAAT 1399  
QY 1310 GATCAAGCTGAATATGACAGCTTTAAGAGCATGTCAGTGGCTTGGAAAGTGTG 1369  
DB 1400 GATCAAGCTGAATATGACAGCTTTAAGAGCATGTCAGTGGCTTGGAAAGTGTG 1459  
QY 1370 GGGATCCAGTTCAGAGAGAGGAGGATGTTGCTCCAGTCTGGGCAAGAGATGCTAT 1429  
DB 1460 GGGATCCAGTTCAGAGAGAGGAGGATGTTGCTCCAGTCTGGGCAAGAGATGCTAT 1519  
QY 1430 GCGGAATTTCTGATATGCAAGAAAGCTCCACCATTTCTTTGATGATGTTTAAAGT 1489  
DB 1520 GCGGAATTTCTGATATGCAAGAAAGCTCCACCATTTCTTTGATGATGTTTAAAGT 1579  
QY 1490 CTCAGTTCCTATATATGAAACAGAGTCTGACAGCTCTGTGCTGATGCTGTC 1549  
DB 1580 CTCAGTTCCTATATATGAAACAGAGTCTGACAGCTCTGTGCTGATGCTGTC 1639  
QY 1550 TGAATATGATGATCTAGGAAAGCAATTTTTTTTCTTGAACCTTAAGCTTATTA 1609  
DB 1640 TGAATATGATGATCTAGGAAAGCAATTTTTTTTCTTGAACCTTAAGCTTATTA 1699  
QY 1610 TTAAGACAGACAGATTTCCATTTTATCATGAGAGATCTTGTGTGATGATGCA 1669  
DB 1700 TTAAGACAGACAGATTTCCATTTTATCATGAGAGATCTTGTGTGATGATGCA 1759  
QY 1670 GGATGAGTGCATCCCTTTAAGAAAGTATATGCTGACCTGCTGCTGCTAATATCTA 1729  
DB 1760 GGATGAGTGCATCCCTTTAAGAAAGTATATGCTGACCTGCTGCTGCTAATATCTA 1818  
QY 1730 ATTTCAGATGCTTTTGTAGATGATGAGTATTTGAGGACCATATTTGGAGTCTAG 1789  
DB 1819 ATTTCAGATGCTTTTGTAGATGATGAGTATTTGAGGACCATATTTGGAGTCTAG 1878  
QY 1790 ATTGAGTGAATGCGAGAAAGGCGCATCTCATTTGAGATATTTAAGTGAACCAACTAG 1849  
DB 1879 ATTGAGTGAATGCGAGAAAGGCGCATCTCATTTGAGATATTTAAGTGAACCAACTAG 1938  
QY 1850 TTTCGGAAATCTACAGAGAGAGGAAATGAGATGAGAGCTGATGATAGGACTTG 1909



Db 140 GCCTTCAGCTAGCGAGCGGCTCGCGGTGAGATCGTCAGCGCTGCATCCATGACAGTCC 199  
 QY 188 TATGAAGGCTTACATCATCACCACCAAGCTTCTGCCCAAGAGCAGAAATCTGCGG 247  
 Db 200 TATGAAGGCTTACATCATCACCACCAAGCTTCTGCCCAAGAGCAGAAATCTGCGG 259  
 QY 248 CACCATGATGATGCTTGTGGATTCCTGTGACCAATTAACAGAGTGTGACTTACAG 307  
 Db 260 CACCATGATGATGCTTGTGGATTCCTGTGACCAATTAACAGAGTGTGACTTACAG 319  
 QY 308 AATGAGCAACTGCTCTGATTAAGATATATTTGCCGAGACAAATTCCTATTGTGTG 367  
 Db 320 AATGAGCAACTGCTCTGATTAAGATATATTTGCCGAGACAAATTCCTATTGTGTG 379  
 QY 368 GAGGAGCAACTATTAATCATTTGAATCTGCTGCGAAATTCCTGTCATTAACCAAGCC 427  
 Db 380 GAGGAGCAACTATTAATCATTTGAATCTGCTGCGAAATTCCTGTCATTAACCAAGCC 439  
 QY 428 CAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAAGCTTGAAGAGAGATGT 487  
 Db 440 CAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAAGCTTGAAGAGAGATGT 499  
 QY 488 CTGTACTTCAAAAGCGCTTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCTGCATCA 547  
 Db 500 CTGTACTTCAAAAGCGCTTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCTGCATCA 559  
 QY 548 CATGACAAAGCAAGGCGCAGGAGCTTGCAGATTTTAAAGAAAGAGAAATCTCAT 607  
 Db 560 CATGACAAAGCAAGGCGCAGGAGCTTGCAGATTTTAAAGAAAGAGAAATCTCAT 619  
 QY 608 AGTGAATTTCTCATGCTCAACATACGAGAAAGGTGTGCTCCCTTGGAGTCTCTG 667  
 Db 620 AGTGAATTTCTCATGCTCAACATACGAGAAAGGTGTGCTCCCTTGGAGTCTCTG 679  
 QY 668 AAGTCTCTAACCTTGCATCTCTTGGCTTCATGCTGACAGGAGTCTTAAGATGAGC 727  
 Db 680 AAGTCTCTAACCTTGCATCTCTTGGCTTCATGCTGACAGGAGTCTTAAGATGAGC 739  
 QY 728 TTGGATTAAGAGGAGTGAATGATCTGCTGCTGGCTCTTGGAGAACTAAGAGATTT 787  
 Db 740 TTGGATTAAGAGGAGTGAATGATCTGCTGCTGGCTCTTGGAGAACTAAGAGATTT 799  
 QY 788 CACAGAGCTATAATCAGAAAGATGTTTCGAAAAATAGCAGAGCTATCAACATGATC 847  
 Db 800 CACAGAGCTATAATCAGAAAGATGTTTCGAAAAATAGCAGAGCTATCAACATGATC 859  
 QY 848 TTCCATCAATTTGGCTTCAAGAAATTCACGAGTACTGATCAGAGGAGAAATGACACA 907  
 Db 860 TTCCATCAATTTGGCTTCAAGAAATTCACGAGTACTGATCAGAGGAGAAATGACACA 919  
 QY 908 CTGGAGCTAGTAACCGCTTCTAAGAAG----- 938  
 Db 920 CTGGAGCTAGTAACCGCTTCTAAGAAGAGTATTAAGAGCTCTGAACAAGTAACTAAG 979  
 QY 939 -----GACTGTGCTCC 949  
 Db 980 AGATATCCCGGAAACAAACCGATGGGTTAAACCGCTTTTGGACAGAGCTGTGCTCC 1039  
 QY 950 ATTGTCCCGCTGTCTATGCTTGAAGTATCTGATCTTCTGAAGTGGAGAGATCTGTT 1009  
 Db 1040 ATTGTCCCGCTGTCTATGCTTGAAGTATCTGATCTTCTGAAGTGGAGAGATCTGTT 1099  
 QY 1010 CTGGAACCTGCTCTGTAATTCGTCAAGATTTCTCCAGGGCCCAAGCCCTACAGCCACT 1069  
 Db 1100 CTGGAACCTGCTCTGTAATTCGTCAAGATTTCTCCAGGGCCCAAGCCCTACAGCCACT 1159  
 QY 1070 CCAATAAAGATGCATCAATGAAGTGAAGACAGAGAGTATTAACCTGTGAGCTC 1129  
 Db 1160 CCAATAAAGATGCATCAATGAAGTGAAGACAGAGAGTATTAACCTGTGAGCTC 1219  
 QY 1130 TGTGATGAATCATCATTTGGGATCGCGAAATGGGACGCAATAAATCCAAATCCAC 1189

Db 1220 TGTGATGAATCATCATTTGGGATTCGGAAATGGGACGCGACATAAATCCAAATCCAC 1279  
 QY 1190 TTGAACCACTGAAGAAAGAGAGATGAGTGCATGATGCTGCAACACCATTAAGAAAT 1249  
 Db 1280 TTGAACCACTGAAGAAAGAGAGATGAGTGCATGATGCTGCAACACCATTAAGAAAT 1339  
 QY 1250 CAGAGTGTTCCTCCAGACTATTAACAAAGAACTAAAGGAAAGGATCCCGAGGAGAAAT 1309  
 Db 1340 CAGAGTGTTCCTCCAGACTATTAACAAAGAACTAAAGGAAAGGATCCCGAGGAGAAAT 1399  
 QY 1310 GATCAAGAGCTGAATTCACAGCTTTAAGAGACATGTCCATGAGCTTGGAAAGGTGTG 1369  
 Db 1400 GATCAAGAGCTGAATTCACAGCTTTAAGAGACATGTCCATGAGCTTGGAAAGGTGTG 1459  
 QY 1370 GGGATCCAGTTCAGAGAGGAGGATGTTTGTCTCCAGCTGAGGCAAGAGAGTGTAT 1429  
 Db 1460 GGGATCCAGTTCAGAGAGGAGGATGTTTGTCTCCAGCTGAGGCAAGAGAGTGTAT 1519  
 QY 1430 GCGGAATTCCTGCTATGACAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGT 1489  
 Db 1520 GCGGAATTCCTGCTATGACAGAAAGCTCCACCATTTCTTTGATGTGTTTAAAGT 1579  
 QY 1490 CTCACGTTCTCTATTAATAGAAACAGAGTCTGTACGCTCTTGTGCTGATGTCTC 1549  
 Db 1580 CTCACGTTCTCTATTAATAGAAACAGAGTCTGTACGCTCTTGTGCTGATGTCTC 1639  
 QY 1550 TGGAAATGATGATGCTCAGAGAAAGCATTTTCTTTTGAACCTTAAGAGTCTATTA 1609  
 Db 1640 TGGAAATGATGATGCTCAGAGAAAGCATTTTCTTTTGAACCTTAAGAGTCTATTA 1659  
 QY 1610 TTAAGACAGACAGATTCACATTTTATACATGAGAGATCTTCTTGTGAGTAATCA 1669  
 Db 1700 TTAAGACAGACAGATTCACATTTTATACATGAGAGATCTTCTTGTGAGTAATCA 1759  
 QY 1670 GGATGATGCTGATCCCTTTAAAGAAATTTATGCTCCCTGACCTGCTGCTAAATATCTA 1729  
 Db 1760 GGATGATGCTGATCCCTTTAAAGAAATTTATGCTCCCTGACCTGCTGCTAAATATCTA 1818  
 QY 1730 ATTCCAGATGCTTTTGTAGATGACGTAAGATTTGTGACCCACATATTTGGAGATCTAG 1789  
 Db 1819 ATTCCAGATGCTTTTGTAGATGACGTAAGATTTGTGACCCACATATTTGGAGATCTAG 1878  
 QY 1790 ATTGAGTGAATGCGAGAAAGGCGCATCTCCATGAGATGATTAAGTGAACCAACTAG 1849  
 Db 1879 ATTGAGTGAATGCGAGAAAGGCGCATCTCCATGAGATGATTAAGTGAACCAACTAG 1938  
 QY 1850 TTCTCGAATTCCTACAGAAAGAGGGAATCAGACTGAGAAAGCTGTGACATGAGACTTG 1909  
 Db 1939 TTCTCGAATTCCTACAGAAAGAGGGAATCAGACTGAGAAAGCTGTGACATGAGACTTG 1998  
 QY 1910 AAGACCAAGACTTTGAAATTTGCGAGCTGCTCATGTGTGATTAATTCACCTGCTCT 1969  
 Db 1999 AAGACCAAGACTTTGAAATTTGCGAGCTGCTCATGTGTGATTAATTCACCTGCTCT 2058  
 QY 1970 TTCTATGAGTTACAAATCTATATTTTATGAAGTTTAAATAAGAAAAATTTACAG 2029  
 Db 2059 TTCTATGAGTTACAAATCTATATTTTATGAAGTTTAAATAAGAAAAATTTACAG 2118  
 QY 2030 AAAAAAAAAA 2041  
 Db 2119 AAAAAAAAAA 2130

RESULT 2  
 US-10-380-731-119  
 ; Sequence 119, Application US/10380731  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 21272-114  
 ; CURRENT APPLICATION NUMBER: US/10/380, 731  
 ; CURRENT FILING DATE: 2003-03-14  
 ; PRIOR APPLICATION NUMBER: 09/659, 671



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:37:18 ; Search time 1151.24 Seconds  
(without alignments)  
7947.624 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041  
Sequence: 1 CTGCCTAAGATGGCGTCCG.....TTTACAGAAAAA 2041

Scoring table: IDENTITY\_NUC  
Gap 10.0, Gapext 1.0

Searched: 593009 seqs, 224144623 residues

Total number of hits satisfying chosen parameters: 11860018

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending Patents\_NA.New:\*  
1: /cgn2\_6/ptodata/2/pna/PCT\_NEW.COMB.seq:\*  
2: /cgn2\_6/ptodata/2/pna/US06\_NEW.COMB.seq:\*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW.COMB.seq:\*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW.COMB.seq:\*  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW.COMB.seq:\*  
6: /cgn2\_6/ptodata/2/pna/US09\_NEW.COMB.seq2:\*  
7: /cgn2\_6/ptodata/2/pna/US09\_NEW.COMB.seq3:\*  
8: /cgn2\_6/ptodata/2/pna/US10\_NEW.COMB.seq:\*  
9: /cgn2\_6/ptodata/2/pna/US10\_NEW.COMB.seq2:\*  
10: /cgn2\_6/ptodata/2/pna/US60\_NEW.COMB.seq:\*  
11: /cgn2\_6/ptodata/2/pna/US60\_NEW.COMB.seq2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1921.2	94.1	2130	6	US-09-659-671A-134 Sequence 134, App
2	1921.2	94.1	2130	8	US-10-380-731-119 Sequence 119, App
3	1869.6	91.6	2119	8	US-10-170-235-40554 Sequence 40554, A
4	1869.6	91.6	2119	11	US-60-453-135-275 Sequence 275, App
5	1869.6	91.6	2119	11	US-60-453-135-275 Sequence 275, App
6	1556.8	76.3	1749	9	US-10-342-887-1692 Sequence 1692, App
7	1354.6	66.4	1719	11	US-10-170-235-41123 Sequence 41123, A
8	1353.8	66.3	1719	11	US-60-453-135-274 Sequence 274, App
9	1353.8	66.3	1719	11	US-60-453-135-274 Sequence 274, App
10	1073.8	52.6	1271	8	US-10-170-235-41124 Sequence 41124, A
11	1073.8	52.6	1271	11	US-60-453-135-276 Sequence 276, App
12	1073.8	52.6	1271	11	US-60-453-135-276 Sequence 276, App
13	1051.2	51.5	2103	9	US-10-144-771-2119 Sequence 2119, App
14	540.8	26.5	555	6	US-09-532-315B-3472 Sequence 3472, App
15	375.4	18.4	392	6	US-09-532-315B-3471 Sequence 3471, App
16	349.2	17.1	500	6	US-09-912-293-11181 Sequence 11181, A
17	320	15.3	334	6	US-09-912-293-109804 Sequence 109804, A
18	291.4	14.3	293	6	US-09-532-315B-9521 Sequence 9521, App
19	285.8	14.0	289	6	US-09-532-315B-9519 Sequence 9519, App
20	278.2	13.6	434	6	US-09-912-293-54990 Sequence 54990, A
21	264	12.9	264	6	US-09-532-315B-9514 Sequence 9514, App
22	261.8	12.8	265	6	US-09-532-315B-9518 Sequence 9518, App

23	261.4	12.8	292	6	US-09-532-315B-9506 Sequence 9506, App
24	257.8	12.6	253	6	US-09-532-315B-9510 Sequence 9510, App
25	256	12.5	256	6	US-09-532-315B-9508 Sequence 9508, App
26	255	12.5	255	6	US-09-532-315B-9511 Sequence 9511, App
27	247	12.1	247	6	US-09-532-315B-9505 Sequence 9505, App
28	242.4	11.9	245	6	US-09-532-315B-9523 Sequence 9523, App
29	239.2	11.7	287	6	US-09-532-315B-9504 Sequence 9504, App
30	239	11.7	240	6	US-09-532-315B-9515 Sequence 9515, App
31	237.2	11.6	266	6	US-09-532-315B-9507 Sequence 9507, App
32	235.8	11.6	246	9	US-10-349-781-40576 Sequence 40576, A
33	232	11.4	255	6	US-09-532-315B-9520 Sequence 9520, App
34	230	11.3	235	6	US-09-532-315B-9513 Sequence 9513, App
35	230	11.3	226	6	US-09-532-315B-9525 Sequence 9525, App
36	227.4	11.1	254	6	US-09-532-315B-9516 Sequence 9516, App
37	220.4	10.8	498	6	US-09-912-293-203132 Sequence 203132, App
38	208.2	10.2	288	6	US-09-532-315B-7599 Sequence 7599, App
39	200.6	9.8	201	11	US-60-453-135-18354 Sequence 18354, A
40	200.6	9.8	201	11	US-60-453-135-18357 Sequence 18357, A
41	200.6	9.8	201	11	US-60-453-135-18362 Sequence 18362, A
42	200.6	9.8	201	11	US-60-453-135-18364 Sequence 18364, A
43	200.6	9.8	201	11	US-60-453-135-18368 Sequence 18368, A
44	200.6	9.8	201	11	US-60-453-050-18354 Sequence 18354, A
45	200.6	9.8	201	11	US-60-453-050-18357 Sequence 18357, A

# ALIGNMENTS

```

RESULT 1
US-09-659-671A-134
? Sequence 134, Application US/09659671A
? GENERAL INFORMATION:
? APPLICANT: Tang, Y. Tom
? APPLICANT: Liu, Chenghua
? APPLICANT: Zhou, Ping
? APPLICANT: Asundi, Vinod
? APPLICANT: Wang, Jian-Rui
? APPLICANT: Zhang, Jie
? APPLICANT: Zhao, Qing A.
? APPLICANT: Ren, Feiyun
? APPLICANT: Xue, Aiding J.
? APPLICANT: Yang, Yonghong
? APPLICANT: Wehrman, Tom
? APPLICANT: Drmanac, Radoje T.
? TITLE OF INVENTION: Novel Nucleic Acids and
? FILE REFERENCE: 794
? CURRENT APPLICATION NUMBER: US/09/659,671A
? CURRENT FILING DATE: 2000-09-11
? NUMBER OF SEQ ID NOS: 475
? SOFTWARE: PL-Fl-genes Version 2.0
? SEQ ID NO 134
? LENGTH: 2130
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (191)..(1426)
? US-09-659-671A-134

Query Match          94.1% Score 1921.2; DB 6; Length 2130;
Best Local Similarly 95.9% Pred. No. 0;
Matches 2025; Conservative 0; Mismatches 8; Indels 79; Gaps 2;

```

OY 8 AAGTGGGCGTCCGCGGCGTGCACGACCTTCTGTGGGCTCAGGGGCTG 67  
 DB 20 ACGATTCCGTCGTCGGGCGCTGCACGACCTTCCGTCGACATGGGCGCTG 79  
 OY 68 CAACGACCTTACTCTTGTAGTATCTTCGCGGCGACGCGGCAATCCACGCTG 127  
 DB 80 CAACGACCTTACTCTTGTAGTATCTTCGCGGCGACGCGGCAATCCACGCTG 139  
 OY 128 GCGTTGACGCTAGGCGACGCGGCTCGGCGGTGAGATCGTCACGCTGACTCATCAGGTC 187

QY 1219 GGACTCAGATGCTGTCAACACCATAGAAAAGTCAGAGTGTTCCTCCAGACTATATACAAAGA 1278  
|||||  
Db 413 GGACTCAGATGCTGTCAACACCATAGAAAAGTCAGAGTGTTCCTCCAGACTATACAAAGA 472  
QY 1279 ACCTAAAGGGAAGGATCCCGAGGAGAAATGATCAAGAGTGAATGCAGCGTTTAAAGA 1338  
|||||  
Db 473 ACCTAAAGGGAAGGATCCCGAGGAGAAATGATCAAGAGTGAATGCAGCGTTTAAAGA 532  
QY 1339 GACATGTCCAGTGGCCTTTGGAAAGGTGGGGATCCAGTTCAGAGGAGGGGTATGT 1398  
|||||  
Db 533 GACATGTCCAGTGGCCTTTGGAAAGGTGGGGATCCAGTTCAGAGGAGGGGTATGT 592  
QY 1399 TTGCTCCCGAGTGGGCAAGAGAGTGTATGGGAAATCTCGCAATAGCAGAAAAGCTC 1458  
|||||  
Db 593 TTGCTCCCGAGTGGGCAAGAGAGTGTATGGGAAATCTCGCAATAGCAGAAAAGCTC 652  
QY 1459 CCACATTTCTTTTGTATGTGTTTAAAGTCTCAGCTTCTATATAGAAACAGCAGG 1518  
|||||  
Db 653 CCACATTTCTTTTGTATGTGTTTAAAGTCTCAGCTTCTATATAGAAACAGCAGG 712  
QY 1519 TCTTGACGCTCCTTGTTGTGTGATGTGTCTGGAATGATGTAGTTCAGGAAAGCA -TT 1577  
|||||  
Db 713 TCTTGACGCTCCTTGTTGTGTGATGTGTCTGGAATGATGTAGTTCAGGAAAGCACTT 772  
QY 1578 TTTTTCCTTTGAACCTTAAAGTCTATATATTAAGCAGACAGATTCACATTTT 1637  
|||||  
Db 773 TTTTTCCTTTGAACCTTAAAGTCTATATATTAAGCAGACAGATTCACATTTT 832  
QY 1638 ATACATGAGAGATCTCTTTGTGTGAATACACAGATGATGCATCCCTTTAAAGAGT 1697  
|||||  
Db 833 ATACATGAGAGATCTCTTTGTGTGAATACACAGATGATGCATCCCTTT -AAAGAGT 891  
QY 1698 TTTATGTCCTGACTGCTGCTAAATATCTAAATTCAGAGTGTCTTTGTAGATGACTGA 1757  
|||||  
Db 892 TTTATGTCCTGACTGCTGCTAAATATCTAAATTCAGAGTGTCTTTGTAGATGACTGA 951  
QY 1758 AGTATTTGTGAGCCACATATTTGGAGTTCAGATTTGAGTAATGGCAGAAAGGCCAT 1817  
|||||  
Db 952 AGTATTTGTGAGCCACATATTTGGAGTTCAGATTTGAGTAATGGCAGAAAGGCCAT 1011  
QY 1818 CTCATTTGAGATGATTAAGTGAACCAAACTAGTTCGGAATTCACAGAGAAAGAGGA 1877  
|||||  
Db 1012 CTCATTTGAGATGATTAAGTGAACCAAACTAGTTCGGAATTCACAGAGAAAGAGGA 1071  
QY 1878 ATCAGACTGAGGAAGTGTGACATAGGACTTGAAGACCAAGACTTGAATTTGGAGC 1937  
|||||  
Db 1072 ATCAGACTGAGGAAGTGTGACATAGGACTTGAAGACCAAGACTTGAATTTGGAGC 1131  
QY 1938 TGCATGCTGATTTATATCACTGCTGCTTCTATGAGTTACAAATCTATATTTT 1997  
|||||  
Db 1132 TGCATGCTGATTTATATCACTGCTGCTTCTATGAGTTACAAATCTATATTTT 1191  
QY 1998 ATTGAAGTTTAAATTAAGAAAAATTTACAAG 2029  
|||||  
Db 1192 ATTGAAGTTTAAATTAAGAAAAATTTACAAG 1223

Search completed: April 21, 2003, 17:14:09  
Job time : 6044.18 secs





QY	956	CCCCGCTGATATGGCTTTAGAGGATATGATGTCTGAGTGGGAGAGACTGCTCTGAA	1015
Db	661	CCCCCTGTCTAGGGCTTAGAGGGATCTGATGTCTGAAAGTGGGAAGACTGTGCTTTGAA	720
QY	1016	CGTGCTCTTAAATCTGCGAAAGTTTCATCCAGGGGCGCAAAACCCTACAGCCACTCCATA	1075
Db	721	CGTGCTCTTAAATCTGCGAAAGTTTCATCCAGGGGCGCAAAACCCTACAGCCACTCCATA	780
QY	1076	AAAGTGCATACATGAGCTGAGAACAAAGAAAGTTATCACCTGTGAGCTCTGTGAT	1135
Db	781	AAAGTGCATACATGAGCTGAGAACAAAGAAAGTTATCACCTGTGAGCTCTGTGAT	840
QY	1136	CGAATCATCATTTGGGGATCGCGAATGGGCGAGCATPAAATCCAAATCCCACTTGAC	1195
Db	841	CGAATCATCATTTGGGGATCGCGAATGGGCGAGCATPAAATCCCAATCCCACTTGAC	900
QY	1196	CAACTGAGAAAAAGAAAGAGATTGGACTCAGATGCTGTGCACACCATGAAAGTCAGAT	1255
Db	901	CAACTGAGAAAAAGAAAGAGATTGGACTCAGATGCTGTGCACACCATGAAAGTCAGAT	960
QY	1256	GTTTCCCCACACATATAACAAAGAAAGCTAAGGAAGGGATCCCGAGGGCGAAGATGATAA	1315
Db	961	GTTTCCCCACACATATAACAAAGAAAGCTAAGGAAGGGATCCCGAGGGCGAAGATGATAA	1020
QY	1316	GAGCTGAATATGACAGCGTTTAAAGAGACATGTCCAGTGGCCCTTTGGAAGAGTGGTGGGATC	1375
Db	1021	GAGCTGAATATGACAGCGTTTAAAGAGACATGTCCAGTGGCCCTTTGGAAGAGTGGTGGGATC	1080
QY	1376	CAGTTCCAGAGAGGAGGGGTATGTTGTCTCCACAGTGTGGGCAAGAGAGTCTATCCGAA	1435
Db	1081	CAGTTCCAGAGAGGAGGGGTATGTTGTCTCCACAGTGTGGGCAAGAGAGTCTATCCGAA	1140
QY	1436	TTCTCTGATACAGAGAAAGCTCCCAACATTTCTTTGATGTGGTTTTAAAGTCTCAG	1495
Db	1141	TTCTCTGATACAGAGAAAGCTCCCAACATTTCTTTGATGTGGTTTTAAAGTCTCAG	1200
QY	1496	TTCTCTTAATAGAAACAGCAGGTCTTGTACGCTCTGTGTGCTGATGTCTGAAA	1555
Db	1201	TTCTCTTAATAGAAACAGCAGGTCTTGTACGCTCTGTGTGCTGATGTCTGAAA	1260
QY	1556	TGATGTAGTCAGAGAAAGCAATTTTTTTTTCTTTGAACTTAAAGTTCTATATTAAA	1615
Db	1261	TGATGTAGTCAGAGAAAGCAATTTTTTTTTCTTTGAACTTAAAGTTCTATATTAAA	1320
QY	1616	GCAGCAGACATCCACATTTTTTATACATGAGATCTCTTGTGTGAATACAGAGATTG	1675
Db	1321	GCAGCAGACATCCACATTTTTTATACATGAGATCTCTTGTGTGAATACAGAGATTG	1380
QY	1676	ACTGCATCCCTTAAAGAAGTTTATGTGCCCTGACTGTGGCTAAATATCTAATTTCC	1735
Db	1381	ACTGCATCCCTTAAAGAAGTTTATGTGCCCTGACTGTGGCTAAATATCTAATTTCC	1439
QY	1736	AGATGCTTTTGTAGATACAGAGATTTTGTGAGCCACATATTGGAGTTCTAGATTGGA	1795
Db	1440	AGATGCTTTTGTAGATACAGAGATTTTGTGAGCCACATATTGGAGTTCTAGATTGGA	1499
QY	1796	GTGAATGGCAGAGAAAGGCCATCTCCATTTGAGATGATTTAAGTGAACCAACTAGTTCTCG	1855
Db	1500	GTGAATGGCAGAGAAAGGCCATCTCCATTTGAGATGATTTAAGTGAACCAACTAGTTCTCA	1559
QY	1856	GAATTCACAGAGAAAGGAGGATCGACTGAGGAAAGCTGTACATAGGACTTGAGAGCC	1915
Db	1560	GAATTCACAGAGAAAGGAGGATCGACTGAGGAAAGCTGTACATAGGACTTGAGAGCC	1619
QY	1916	AAAGACTTTGAATTTTGCAGAGCTCATGTGTGAATTTATTCACCTGCTGCTTTCTAT	1975
Db	1620	AAAGACTTTGAATTTTGCAGAGCTCATGTGTGAATTTATTCACCTGCTGCTTTCTAT	1679
QY	1976	TGAATTCACCAATCTATTTTTTTATGAGTTTAAATAAAGAAAAATTTACAGAGAAAA	2035
Db	1680	TGAATTCACCAATCTATTTTTTTATGAGTTTAAATAAAGAAAAATTTAAAGAAAAAA	1739

DB	Seq ID	Seq	Score	DB	Length
Db	1740	AAAAAA 1745			
Qy	2036	AAAAAA 2041			
Qy	12	US-09-757-028-874			
		Sequence 874, Application US/09757028			
		GENERAL INFORMATION:			
		APPLICANT: Rosen et al.			
		TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies			
		FILE REFERENCE: PM001			
		CURRENT APPLICATION NUMBER: US/09/757,028			
		CURRENT FILING DATE: 2001-01-09			
		PRIOR APPLICATION NUMBER: 60/179,065			
		PRIOR FILING DATE: 2000-01-31			
		PRIOR APPLICATION NUMBER: 60/180,628			
		PRIOR FILING DATE: 2000-02-04			
		NUMBER OF SEQ ID NOS: 2660			
		SOFTWARE: PatentIn Ver. 2.0			
		SEQ ID NO 874			
		LENGTH: 1183			
		TYPE: DNA			
		ORGANISM: Homo sapiens			
		FEATURE:			
		NAME/KEY: SITE			
		LOCATION: (1172)			
		OTHER INFORMATION: n equals a,t,g, or c			
		NAME/KEY: SITE			
		LOCATION: (1177)			
		OTHER INFORMATION: n equals a,t,g, or c			
		US-09-757-028-874			
		Query Match	53.3%	Score 1088.2;	DB 29; Length 1183;
		Best Local Similarity	98.6%	Pred. No. 1.4e-226;	
		Matches 1107; Conservative 1;	Mismatches 14;	Indels 1;	Gaps 1
Qy	919	TAAACGCTTTAAAGAAAGACCTGGTCCCTATTTGCCCCCGTGTATAGCTAGAGT	978		
Db	15	TAAAAACCGTTTTTTGAGAGACACTGGTCCATGTGCCCCCTGTCTATAGCTTAGAGT	74		
Qy	979	ATCTGATGTCGGAAGTGGAGAGTCTGTTCTTGAACTGCTTTGAATCGTCAAG	1038		
Db	75	ATCTGATGTCGGAAGTGGAGAGTCTGTTCTTGAACTGCTTTGAATCGTCAAG	134		
Qy	1039	TTTCATCCAGGGCCCAACACCTACGACCACTCCCAATTAAGATGCCATACATGAACCTGA	1098		
Db	135	TTTCATCCAGGGCCCAACACCTACGACCACTCCCAATTAAGATGCCATACATGAACCTGA	194		
Qy	1099	GAAACAAGAAATTTATCAGCTGTGTACCTGTGTGATCGATCATCATATTGGGATGGCA	1158		
Db	195	GAAACAAGAAATTTATCAGCTGTGTACCTGTGTGATCGATCATCATATTGGGATGGCA	254		
Qy	1159	ATGGCAGCGCACATTAATAATCCAAATCCCACTTGAACCACTGAGAGAAAGAAAGATT	1218		
Db	255	ATGGCAGCGCACATTAATAATCCAAATCCCACTTGAACCACTGAGAGAAAGAAAGATT	314		
Qy	1219	GGACATCAGTGTGTCAACACCATAGAAAGTCAAGTGTTCCTCCCGACATATAACAAGA	1278		
Db	315	GGACATCAGTGTGTCAACACCATAGAAAGTCAAGTGTTCCTCCCGACATATAACAAGA	374		
Qy	1279	ACCTTAAGGAGAGGATCCCCAGGGCAGATATCAAGAGAGTGAATGACAGGTTTAAGA	1338		
Db	375	ACCTTAAGGAGAGGATCCCCAGGGCAGATATCAAGAGAGTGAATGACAGGTTTAAGA	434		
Qy	1339	GACATGTCACAGTGGCTTTGGAAGGTGGTGGGATCCAGTTCCAGGAGGAGGGGTATGT	1398		
Db	435	GACATGTCACAGTGGCTTTGGAAGGTGGTGGGATCCAGTTCCAGGAGGAGGGGTATGT	494		
Qy	1399	TTGTCTCCCAAGTCTGGGCAAGAGATGCTATGCGGAATTCCTGCTATACAGAAAAGCTC	1458		
Db	495	TTGTCTCCCAAGTCTGGGCAAGAGATGCTATGCGGAATTCCTGCTATACAGAAAAGCTC	554		

Db 721 CCTGCTTGAATGCTGCAAAAGTTTCATCCAGGCGCACAGCTACACCCACTCCATTA 780  
 QY 1076 AAGATGCATACAAAGAGCTGAGAACAGAGATATACCTGTGTACCTCTGTAT 1135  
 Db 781 AAGATGCATACAAAGAGCTGAGAACAGAGATATACCTGTGTACCTCTGTAT 840  
 QY 1136 CGAATCATCATTTGGGGATGCGCAATGGGAGCGCACATATAATTCNAATCCACTGAAAC 1195  
 Db 841 CGAATCATCATTTGGGGATGCGCAATGGGAGCGCACATATAATTCNAATCCACTGAAAC 900  
 QY 1196 CAATGAGAAAAGAGAGATTTGAGCTCAGATGCTGTCACACCATACAGAGAGTCAAGT 1255  
 Db 901 CAATGAGAAAAGAGAGATTTGAGCTCAGATGCTGTCACACCATACAGAGAGTCAAGT 960  
 QY 1256 GTTTCGCCAGACTATAACAAAGCTTAAGAGAGAGGATCCCGAGGGAGAGATGATCA 1315  
 Db 961 GTTTCGCCAGACTATAACAAAGCTTAAGAGAGAGGATCCCGAGGGAGAGATGATCA 1020  
 QY 1316 GAGCTGAATGAGAGGCTTTTAAGAGACATGTCCAGTGGCTTTGGAAGGTGGGGATC 1375  
 Db 1021 GAGCTGAATGAGAGGCTTTTAAGAGACATGTCCAGTGGCTTTGGAAGGTGGGGATC 1080  
 QY 1376 CAGTTGAGAGAGGAGGGATGTGTGTCTCCAGTCTGGGCAAGAGAGTGTATGGGAA 1435  
 Db 1081 CAGTTGAGAGAGGAGGGATGTGTGTCTCCAGTCTGGGCAAGAGAGTGTATGGGAA 1140  
 QY 1436 TTCTCTGCTACAGAAAAGCTCCACCATTTTCTTATGTGGTTTAAAGTCTCAG 1495  
 Db 1141 TTCTCTGCTACAGAAAAGCTCCACCATTTTCTTATGTGGTTTAAAGTCTCAG 1200  
 QY 1496 TTCTCTATATAGAAAACACAGAGTCTGTACGCTCTGTGGGCTGATGTCTGAAA 1555  
 Db 1201 TTCTCTATATAGAAAACACAGAGTCTGTACGCTCTGTGGGCTGATGTCTGAAA 1260  
 QY 1556 TGATGATGTCAGAAAAGCATTTTCTTCTTGAACCTTAAAGGTTCTATATTAAAA 1615  
 Db 1261 TGATGATGTCAGAAAAGCATTTTCTTCTTGAACCTTAAAGGTTCTATATTAAAA 1320  
 QY 1616 GCAGCAGACATTCACATTTTATACATAGAGATCTTCTTGTGGTGAATACAGAGATTG 1675  
 Db 1321 GCAGCAGACATTCACATTTTATACATAGAGATCTTCTTGTGGTGAATACAGAGATTG 1380  
 QY 1676 ACTGCATCCCTTAAAAAGAGTTTATGTCCTGACTGCGCTAAATATCTAATTTCC 1735  
 Db 1381 ACTGCATCCCTT -AAGAGAGTTTATGTCCTGACTGCGCTAAATATCTAATTTCC 1439  
 QY 1736 AGATGCTTTTGTAGATGACTGAGATTTGTGAGCCACATATTGGAGTCTTAGATTGA 1795  
 Db 1440 AGATGCTTTTGTAGATGACTGAGATTTGTGAGCCACATATTGGAGTCTTAGATTGA 1499  
 QY 1796 GTGATGCGAGAAAGGCCATCTCCATTGAGATGATTAAGTGAACCAAACTAGTTCTG 1855  
 Db 1500 GTGATGCGAGAAAGGCCATCTCCATTGAGATGATTAAGTGAACCAAACTAGTTCTG 1559  
 QY 1856 GAATTCACAGAGAGAGAGGAAATCAGACTGAGAGAGCTGACATAGACTGAGAGACC 1915  
 Db 1560 GAATTCACAGAGAGAGAGGAAATCAGACTGAGAGAGCTGACATAGACTGAGAGACC 1619  
 QY 1916 AAGACTTTGAATTTGCGAGCTGCTCATGTGTGAGTATTAATCACTGCTGTTCTAT 1975  
 Db 1620 AAGACTTTGAATTTGCGAGCTGCTCATGTGTGAGTATTAATCACTGCTGTTCTAT 1679  
 QY 1976 TGAGTTACAAATCTATATTTTAAATGAATTTAAATGAAGAAAATTTACAGAGAAAA 2035  
 Db 1680 TGAGTTACAAATCTATATTTTAAATGAATTTAAATGAAGAAAATTTACAGAGAAAA 1739  
 QY 2036 AAAAAA 2041  
 Db 1740 AAAAAA 1745

; Sequence 1692, Application US/10172118  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dai, Hongyue  
 ; APPLICANT: He, Yudong  
 ; APPLICANT: Linsley, Peter  
 ; APPLICANT: Mao, Mao  
 ; APPLICANT: Roberts, Chris  
 ; APPLICANT: Van 't Veer, Laura  
 ; APPLICANT: Van de Vijver, Marc  
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
 ; FILE REFERENCE: 9301-175-999  
 ; CURRENT APPLICATION NUMBER: US/10/172,118  
 ; CURRENT FILING DATE: 2002-06-14  
 ; PRIOR APPLICATION NUMBER: 60/380,770  
 ; PRIOR FILING DATE: 2002-05-14  
 ; NUMBER OF SEQ ID NOS: 2699  
 ; SEQ ID NO 1692  
 ; LENGTH: 1749  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; PUBLICATION INFORMATION:  
 ; DATABASE ACCESSION NUMBER: NM\_017646  
 ; DATABASE ENTRY DATE: 2001-06-18  
 ; US-10-172-118-1692  
 Query Match 76.3%; Score 1556.8; DB 41; Length 1749;  
 Best Local Similarity 95.1%; Pred. No. 0;  
 Matches 1660; Conservative 0; Mismatches 7; Indels 79; Gaps 2;  
 QY 374 ACCAATTATATACATTGAATCTGCTGCGTGGAAAGTCTTGCAATACCAAGCCGAGAG 433  
 Db 1 ACCAATTATATACATTGAATCTGCTGCGTGGAAAGTCTTGCAATACCAAGCCGAGAG 60  
 QY 434 ATGGGCATGAGAAAGTGAATGACCGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 493  
 Db 61 ATGGGCATGAGAAAGTGAATGACCGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 120  
 QY 494 CTTCAACAAAGCCTTAAGCCAGTGGAGCCAGAAATGGCTGCCAAGTGCATCCATGAC 553  
 Db 121 CTTCAACAAAGCCTTAAGCCAGTGGAGCCAGAAATGGCTGCCAAGTGCATCCATGAC 180  
 QY 554 AAACGAAAGTGGCCAGGAGCTTGCAAGTTTGAAGAAAGCAAGATCTCTATAGTGA 613  
 Db 181 AAACGAAAGTGGCCAGGAGCTTGCAAGTTTGAAGAAAGCAAGATCTCTATAGTGA 240  
 QY 614 TTTCTCATCTGCAACATACGGAAGAGTGTGTCCTGAGAGTCTCTGAAATTG 673  
 Db 241 TTTCTCATCTGCAACATACGGAAGAGTGTGTCCTGAGAGTCTCTGAAATTG 300  
 QY 674 TCTAACCTTGCATCTTTGGCTTCAATGCTGACACGAGCAATTCTAGATGAGCGTTGAT 733  
 Db 301 TCTAACCTTGCATCTTTGGCTTCAATGCTGACACGAGCAATTCTAGATGAGCGTTGAT 360  
 QY 734 AAGAGGTGATGACATGCTGCTGCTGGGCTTTGGAGAGCTAAGAAATTTACAGA 793  
 Db 361 AAGAGGTGATGACATGCTGCTGCTGGGCTTTGGAGAGCTAAGAAATTTACAGA 420  
 QY 794 CGCTATATACAGAAATTTTCGAAATATAGCAGAGACTATCAACATGATCTTCCAA 853  
 Db 421 CGCTATATACAGAAATTTTCGAAATATAGCAGAGACTATCAACATGATCTTCCAA 480  
 QY 854 TCAATTTGGCTCAAGAAATTTTCAGAGTACGTGATCAGTGAAGGAAATGACACTGGAG 913  
 Db 481 TCAATTTGGCTCAAGAAATTTTCAGAGTACGTGATCAGTGAAGGAAATGACACTGGAG 540  
 QY 914 ACTAGTAACCACTTTAAAGAAAG----- 939  
 Db 541 ACTAGTAACCACTTTAAAGAAAGTTTGAAGGCTGGAAGCAAGTAAGAGATAT 600  
 QY 940 -----ACCTGCTCCCATTTGTC 955  
 Db 601 GCCCGAAACAAACCGATGGGTTAAAAACCGTTTTTTGAGCAGACCTGTCCTCATTTGTC 660

```

Db      839 AATCAATGGCTTCAAGAAATTCACAGTACCTGATCACTGAGGAAATGCACACTGG 898
QY      912 AGACTAGTACCACTCTTAAAGAAAG----- 938
Db      899 AGACTAGTAAACGAGCTTCTAAAGAAAGTATTGAGGCTCTGAAACAGTAACATAAGAGAT 958
QY      939 -----GACCTGGCTCCCATG 953
Db      959 ATGCCCGGAACAAACCGATGGGTTAAAAACGTTTTTTGAGCAACCTGGTCCCATG 1018
QY      954 TCCCCCTGTCTATGGCTTAGAGTATCTGATGTCTGAAAGTGGAGAGTGTCTTG 1013
Db      1019 TCCCCCTGTCTATGGCTTAGAGTATCTGATGTCTGAAAGTGGAGAGTGTCTTG 1078
QY      1014 AACCTGCTCTTGAATCGTCAAAAGTTTATCCAGGGCCACAAGCCTACAGCCACTCAA 1073
Db      1079 AACCTGCTCTTGAATCGTCAAAAGTTTATCCAGGGCCACAAGCCTACAGCCACTCAA 1138
QY      1074 TAAAGTGGCATACATGAACTGAGACACAAAGAAAGTTATCACTGTGTGACCTGTG 1133
Db      1139 TAAAGTGGCATACATGAACTGAGACACAAAGAAAGTTATCACTGTGTGACCTGTG 1198
QY      1134 ATCGAATCATATGGGGATCGGATGGGAGCGCACATAAATCCAAATCCACTTG 1193
Db      1199 ATCGAATCATATGGGGATCGGATGGGAGCGCACATAAATCCAAATCCACTTG 1258
QY      1194 ACCAATGGAAGAAAGAAAGATTTGACTGATGCTGTCAACACCATAGAAAGTCAG 1253
Db      1259 ACCAATGGAAGAAAGAAAGATTTGACTGATGCTGTCAACACCATAGAAAGTCAG 1318
QY      1254 GTGTTTCCCGAGCTATTAACAAAGAACTTAAGGAGAGATCCCGAGGACAGATGATC 1313
Db      1319 GTGTTTCCCGAGCTATTAACAAAGAACTTAAGGAGAGATCCCGAGGACAGATGATC 1378
QY      1314 AAGAGTGAATGACGCTTTAAGAGACATGTCACATGGCTTTGAAGGTGGTGGGA 1373
Db      1379 AAGAGTGAATGACGCTTTAAGAGACATGTCACATGGCTTTGAAGGTGGTGGGA 1438
QY      1374 TCCAGTTCAAGAGGAGGGATGTTGTCTCCAGTCTGGCCAAAGAGAGTCTATGCG 1433
Db      1439 TCCAGTTCAAGAGGAGGGATGTTGTCTCCAGTCTGGCCAAAGAGAGTCTATGCG 1498
QY      1434 AATTCCTCATAGCAGAAAGCTCCACATTTCTTTGATGTGTTTAAAGTCTCA 1493
Db      1499 AATTCCTCATAGCAGAAAGCTCCACATTTCTTTGATGTGTTTAAAGTCTCA 1558
QY      1494 CGTTCTTAATAGAAACAGCAGTCTGTGCTCCCTGCTGGCTGTGCTGGA 1553
Db      1559 CGTTCTTAATAGAAACAGCAGTCTGTGCTCCCTGCTGGCTGTGCTGGA 1618
QY      1554 AATGATGATGTCAGGAAAGCATTTTTTTCTTTGAACCTTAAGGTTCTATTATTA 1613
Db      1619 AATGATGATGTCAGGAAAGCA-TTTTTTTCTTTGAACCTTAAGGTTCTATTATTA 1677
QY      1614 AAGCAGCAGATTCACATTTTTTATACATGAGATCTTTGTGTGTAATCCAGAT 1673
Db      1678 AAGCAGCAGATTCACATTTTTTATACATGAGATCTTTGTGTGTAATCCAGAT 1737
QY      1674 TGACTGATCCCTTTAAAGAAAGTTTATGCTCCCTGACTGTGCTAAATATATATTT 1733
Db      1738 TGACTGATCCCTTT-AAAGAAAGTTTATGCTCCCTGACTGTGCTAAATATATTT 1796
QY      1734 CCAGATGCTTTGTAGATGACTGAAGTATTGTGAGCCACATATTTGGAGTTCTAGATT 1793
Db      1797 CCAGATGCTTTGTAGATGACTGAAGTATTGTGAGCCACATATTTGGAGTTCTAGATT 1856
QY      1794 GA 1795
Db      1857 GA 1858

```

RESULT 10  
PCT-US02-18947-1692

```

; Sequence 1692, Application PCT/US0218947
; GENERAL INFORMATION:
; APPLICANT: Rosetta Inpharmatics
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-228
; CURRENT APPLICATION NUMBER: PCT/US02/18947
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1692
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_017646
; DATABASE ENTRY DATE: 2001-06-18
PCT-US02-18947-1692

Query Match      76.3%; Score 1556.8; DB 1; Length 1749;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1660; Conservative 0; Mismatches 7; Indels 79; Gaps 2;

QY      374 ACCAATTATTACATTTGAAATCTGCTGTGAAAGTCTTGTCAATACCAAGCCCGAGAG 433
Db      1 ACCAATTATTACATTTGAAATCTGCTGTGAAAGTCTTGTCAATACCAAGCCCGAGAG 60
QY      434 ATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTTGAAGAGAGATGGTCTGTA 493
Db      61 ATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTTGAAGAGAGATGGTCTGTA 120
QY      494 CTTCACAAAGCCCTAAGCCAGAGTGGAGCCGAAATGGCTGCCAAGCTGATCCACATGAC 553
Db      121 CTTCACAAAGCCCTAAGCCAGAGTGGAGCCGAAATGGCTGCCAAGCTGATCCACATGAC 180
QY      554 AAGCGAAAGTGGCCAGAGCTTGCAGATTTTGAAAGAAACAGAAATCTCTCATAGTGA 613
Db      181 AAGCGAAAGTGGCCAGAGCTTGCAGATTTTGAAAGAAACAGAAATCTCTCATAGTGA 240
QY      614 TTCTCCTCATGCTCAACATACAGAAAGTGTGTGCTCCCTGGAGTCTCTGAAGTTG 673
Db      241 TTCTCCTCATGCTCAACATACAGAAAGTGTGTGCTCCCTGGAGTCTCTGAAGTTG 300
QY      674 TCTAACCTTGATCTCTTGGCTTCATGCTGACCGAGCACTTCTAGTAGGCCCTTGAT 733
Db      301 TCTAACCTTGATCTCTTGGCTTCATGCTGACCGAGCACTTCTAGTAGGCCCTTGAT 360
QY      734 AAGAGGTGATGACATGCTGTGCTGTGGGCTCTTGAGAGAACTAAGATTTTCACAGA 793
Db      361 AAGAGGTGATGACATGCTGTGCTGTGGGCTCTTGAGAGAACTAAGATTTTCACAGA 420
QY      794 CGCTATATCAGAAAGATGTTTGGAAATATAGCAGACTATACATGATATCTTCCAA 853
Db      421 CGCTATATCAGAAAGATGTTTGGAAATATAGCAGACTATACATGATATCTTCCAA 480
QY      854 TCAATTGCTTCAAGGAATTTCAAGATACCTGATCACTGAGGAAATACACACTGGAG 913
Db      481 TCAATTGCTTCAAGGAATTTCAAGATACCTGATCACTGAGGAAATACACACTGGAG 540
QY      914 ACTAGTAACAGCTTCTAAGAAAG----- 939
Db      541 ACTAGTAACAGCTTCTAAGAAAGATATTAGAGCTCTGAACAGTAACATAAGAGATAT 600
QY      940 -----ACCTGTGCCATTTGTC 955
Db      601 GCCCGGAACAAACCGATGGGTTAAACCGTTTTTTGAGCAGAGCTGTCATTTGTC 660
QY      956 CCCCTGTCTATGAGCTTAGAGATATCTGATGTCGAAGTGGAGAGTGTGTTTGA 1015
Db      661 CCCCTGTCTATGAGCTTAGAGATATCTGATGTCGAAGTGGAGAGTGTGTTTGA 720
QY      1016 CCGTCTTGAATCGTGAAGTTCATCCAGGGCCACAAGGCTCAGCAGCTCCATA 1075

```

```

Db 1571 TAATAGAAACAGCAGTCTTGTACGCTCTTGTGTGCTGATGTCTGTAATGATGTA 1630
QY 1563 GTTCAGGAAGATTTTTTTTTTTTGTGAACCTTAAGSTTCTATATTAAGACGAC 1622
Db 1631 GTTCAGGAAGATTTTTTTTTTTTGTGAACCTTAAGSTTCTATATTAAGACGAC 1690
QY 1623 AGATTCACATTTTTATACATGAGATCTCTTTGTGGAATACAGGATGACGACAT 1682
Db 1691 AATTCACATTTTTATACATGAGATCTCTTTGTGGAATACAGGATGACGACAT 1750
QY 1683 CCGTTTAAAGATTTTATGTCCCTGACCTGCTGCTAAATATCTAATTTCCAGATGCT 1742
Db 1751 CCGTTT-AAAGAAATTTTGTCCCTGACCTGCTGCTAAATATCTAATTTCCAGATGCT 1809
QY 1743 TTTGTATGATGATGAGATTTTGTGAGCACAATTTGGAGTCTGTGATTTGAGTAAAG 1802
Db 1810 TTTGTATGATGATGAGATTTTGTGAGCACAATTTGGAGTCTGTGATTTGAGTAAAG 1869
QY 1803 GGAGAAAGGCGCATCTCCATGATGATTAAGTAAAGTAAACCAACTAGTCTCGGATTCCT 1862
Db 1870 GGAGAAAGGCGCATCTCCATGATGATTAAGTAAAGTAAACCAACTAGTCTCGGATTCCT 1929
QY 1863 ACAGAGAGAGGAGGAATCAGACTGAGAAAGCTGTGACATAGACCTTGAAGCAAGACT 1922
Db 1930 ACAGAGAGAGGAGGAATCAGACTGAGAAAGCTGTGACATAGACCTTGAAGCAAGACT 1989
QY 1923 TTGAATTTGCGAGCTGCTCATGTGTGATTTATACACTGCTGCTTCTATTTGAGTAA 1982
Db 1990 TTGAATTTGCGAGCTGCTCATGTGTGATTTATACACTGCTGCTTCTATTTGAGTAA 2049
QY 1983 CAAATCTATTTTTTTTATGAAGTTTAAATAAGAAAAATTTTCAAGAAAA 2034
Db 2050 CAAATCTATTTTTTTTATGAAGTTTAAATAAGAAAAATTTTCAAGAAAA 2101

```

# RESULT 9

```

; Sequence 2814, Application US/09471275
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; TITLE OF INVENTION: Novel Contigs Obtained
; TITLE OF INVENTION: From Various Libraries
; FILE REFERENCE: 782
; CURRENT APPLICATION NUMBER: US/09/471, 275
; CURRENT FILING DATE: 1999-12-23
; EARLIER APPLICATION NUMBER: US 09/235, 076
; EARLIER FILING DATE: 1999-01-20
; EARLIER APPLICATION NUMBER: US 09/234, 611
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: US 09/240, 371
; EARLIER FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: US 09/277, 227
; EARLIER FILING DATE: 1999-03-25
; EARLIER APPLICATION NUMBER: US 09/271, 490
; EARLIER FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: US 09/293, 972
; EARLIER FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 09/274, 861
; EARLIER FILING DATE: 1999-03-23
; EARLIER APPLICATION NUMBER: US 60/125, 453
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: US 60/126, 605
; EARLIER FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: US 09/306, 350
; EARLIER FILING DATE: 1999-05-07
; EARLIER APPLICATION NUMBER: US 09/399, 720
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: US 09/404, 284
; EARLIER FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER:
; EARLIER FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 10451
; SOFTWARE: pt_CT_genes Version 1.0

```

```

; SEQ ID NO 2814
; LENGTH: 1858
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1843)...(1283)
; OTHER INFORMATION: similar to g11419759 in the genpept database release 114,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-471-275-2814

```

```

Query Match 80.6%; Score 1646; DB 18; Length 1858;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1773; Conservative 0; Mismatches 5; Indels 84; Gaps 4;

```

```

QY 12 TGGCGTCCGCGGCGCTGACAGACAGTCTCTGTGGGAGTGCGCTCAGGGGCTGCAAC 71
Db 3 TCGTGTCCTGGCGGCTGACAGAGTCTCTGTGGGAGTGCGCTCAGGGGCTGCAAC 62
QY 72 GGACCTACCTCTTGTAGTATTCGCGGGGCGGCGGCGGCAATCCAGCTGGCGT 131
Db 63 GGACCTACCTCTTGTAGTATTCGCGGGGCGGCGGCGGCAATCCAGCTGGCGT 122
QY 132 TCGAGCTAGCGGCGGCTGCGGCGGAGATGTCAGCGCTGATCCATGACAGTCTATG 191
Db 123 TCGAGCTAGCGGCGGCTGCGGCGGAGATGTCAGCGCTGATCCATGACAGTCTATG 182
QY 192 AAGGCTACATCATCATCACCACAGAGTTTCTGCGGCAAGAGACAGATCTCGGCGAC 251
Db 183 AAGGCTACATCATCATCACCACAGAGTTTCTGCGGCAAGAGACAGATCTCGGCGAC 242
QY 252 ACATGATCAGCTTGTGATTCCTGTGACCAATTCACAGTGGTGGAGTTCAGAAATA 311
Db 243 ACATGATCAGCTTGTGATTCCTGTGACCAATTCACAGTGGTGGAGTTCAGAAATA 302
QY 312 GAGCAACTGCTGATTTGAAGATATATTTCCCGAGACAAAATTCCTATTTGTGGAG 371
Db 303 GAGCAACTGCTGATTTGAAGATATATTTCCCGAGACAAAATTCCTATTTGTGGAG 362
QY 372 GAACCAATTTATCATTTGATGATCTCTGCTGGAAGTTTCTTCATACCAAGCCCGAG 431
Db 363 GAACCAATTTATCATTTGATGATCTCTGCTGGAAGTTTCTTCATACCAAGCCCGAG 422
QY 432 AGATGGGCACTAGAAAGATTTGACCGAAAGTGAGGCTTGAAGAGAGATGGGCTTG 491
Db 423 AGATGGGCACTAGAAAGATTTGACCGAAAGTGAGGCTTGAAGAGAGATGGGCTTG 482
QY 492 TACTTACAAAGCCCTTAACCCAGAGTGAGCCAGAAATGCTGCCAAGCTGCATCCACATG 551
Db 483 TACTTACAAAGCCCTTAACCCAGAGTGAGCCAGAAATGCTGCCAAGCTGCATCCACATG 542
QY 552 ACNAAAGCAAGAGTGCGCAGAGCTTGCAGATTTTGAAGAAACAGGAATCTCTATAGTG 611
Db 543 ACNAAAGCAAGAGTGCGCAGAGCTTGCAGATTTTGAAGAAACAGGAATCTCTATAGTG 602
QY 612 AATTTCTCATGCTCAACATACGGAAGAGTGCTGCTGAGAGTCTCTGAGT 671
Db 603 AATTTCTCATGCTCAACATACGGAAGAGTGCTGCTGAGAGTCTCTGAGT 662
QY 672 TCTCTAACCTTGACATCTTGGCTTCATGCTGACAGCAGGCTGATGATGAGGCTTGG 731
Db 663 TCTCTAACCTTGACATCTTGGCTTCATGCTGACAGCAGGCTGATGATGAGGCTTGG 718
QY 732 ATTAGAGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
Db 719 ATTAGAGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778
QY 792 GACGCTATATCAGAAAGATTTTGGAAATACGAGAGCTATCAACATGATCTTCC 851
Db 779 GACGCTATATCAGAAAGATTTTGGAAATACGAGAGCTATCAACATGATCTTCC 838
QY 852 AATCAATGGCTTCAAGGAATTTCAAGAGTACTGATCTGAGGAGAAATCACAGCTGG 911

```





OTHER INFORMATION: Incyte ID No: 977391.2  
NAME/KEY: unsure  
LOCATION: 2106  
OTHER INFORMATION: a, t, c, g, or other  
US-60-324-185-33690

Query Match 93.5%; Score 1909; DB 76; Length 2109;  
Best Local Similarity 95.9%; Pred. No. 0;  
Matches 2022; Conservative 0; Mismatches 6; Indels 80; Gaps 3;

QY 6 AATAAGATGCGTCCGT-GGCGGCTGCACGAGCAGTTCGTGTGGGAGTGGGCTCAGAGGC 64  
DB 3 AAAAGATGGCGTCCGTGGGGGCTGCACGAGCAGTTCGTGTGGGAGTGGGCTCAGAGGC 62  
QY 65 CTGCAACGAGCCTTACTCTTGTAGTATTCCTGGGGCCACGGGACCGGCAATTCACG 124  
DB 63 CTGCAACGAGCCTTACTCTTGTAGTATTCCTGGGGCCACGGGACCGGCAATTCACG 122  
QY 125 CTGGGCTGTGAGTGCACGAGGCGTGGGGGTGAGTGTGAGCGCTGACCTGATGACG 184  
DB 123 CTGGGCTGTGAGTGCACGAGGCGTGGGGGTGAGTGTGAGCGCTGACCTGATGACG 182  
QY 185 GTCTATGAAGGCTTAGACATCATCACCAACAGGTTTCCGCAAGAGCAGAGATCTGC 244  
DB 183 GTCTATGAAGGCTTAGACATCATCACCAACAGGTTTCCGCAAGAGCAGAGATCTGC 242  
QY 245 CGGCACCATGATGATCAGCTTTGTGATTCCTCTTGTGACCAATTACAGTGTGACTTC 304  
DB 243 CGGCACCATGATGATCAGCTTTGTGATTCCTCTTGTGACCAATTACAGTGTGACTTC 302  
QY 305 AGAATAGAGCAACCTGCTGATGAAATATTTTCCCGAGACAAATTCCTATTGT 364  
DB 303 AGAATAGAGCAACCTGCTGATGAAATATTTTCCCGAGACAAATTCCTATTGT 362  
QY 365 GTGGGAGAACCAATTATTACATGATCTCTGCTGAAATTCCTTGTCAATACCAAG 424  
DB 363 GTGGGAGAACCAATTATTACATGATCTCTGCTGAAATTCCTTGTCAATACCAAG 422  
QY 425 CCCGAGAGATGGGACTGAGAAAGTGTATGACCGAAAGTGGAGCTTGAAGAGAGAT 484  
DB 423 CCCGAGAGATGGGACTGAGAAAGTGTATGACCGAAAGTGGAGCTTGAAGAGAGAT 482  
QY 485 GGTCTGTACTTACAAACCCCTAAGCAGGTGGAGCCAGAAATGGTGGCAAGCTGAT 544  
DB 483 GGTCTGTACTTACAAACCCCTAAGCAGGTGGAGCCAGAAATGGTGGCAAGCTGAT 542  
QY 545 CCACATGACAAACGCAAGTGGCCAGAGGCTTCAAGTTTGAAGAAACAGGATCTCT 604  
DB 543 CCACATGACAAACGCAAGTGGCCAGAGGCTTCAAGTTTGAAGAAACAGGATCTCT 602  
QY 605 CATAGTGAATTTCTCATGCTCAACATACGAGAAAGTGGTGGTCCCTTGGAGTCTCT 664  
DB 603 CATAGTGAATTTCTCATGCTCAACATACGAGAAAGTGGTGGTCCCTTGGAGTCTCT 662  
QY 665 CTGAAGTCTCTAACCTTGCATCTTGGCTTCATGCTGACGAGGAGTTCATGATGAG 724  
DB 663 CTGAAGTCTCTAACCTTGCATCTTGGCTTCATGCTGACGAGGAGTTCATGATGAG 722  
QY 725 CGCTTGAATGAAGGGTGTGATGATGCTGCTGGGCTCTTGAAGAACTAAGAGAT 784  
DB 723 CGCTTGAATGAAGGGTGTGATGATGCTGCTGGGCTCTTGAAGAACTAAGAGAT 782  
QY 785 TTTTCACAGACGCTATTAATCAGAAATGTTTGGAAAAATAGCAGAGCTATCAACATGCT 844  
DB 783 TTTTCACAGACGCTATTAATCAGAAATGTTTGGAAAAATAGCAGAGCTATCAACATGCT 842  
QY 845 ATCTTCAATCAATTTGGCTTCAAGAAATTTCAAGAGTTCACAGAGTTCACAGAGTTC 904  
DB 843 ATCTTCAATCAATTTGGCTTCAAGAAATTTCAAGAGTTCACAGAGTTCACAGAGTTC 902  
QY 905 ACACGTGAGAGCTAGTAACCAAGCTTCTAAAGAAAG----- 938  
DB 903 ACACGTGAGAGCTAGTAACCAAGCTTCTAAAGAAAGGATTTGAGGCTCTGAAGAAAGTACT 902

QY 939 -----GACCTGCT 946  
DB 963 AAGAGATATGCCCGGAACAAAACGATGGGTAAAAACCGTTTGTAGACAGACCTGCT 1022  
QY 947 CCCATTGTCCCCCTGTCTATAGGCTTATAGAGTATCTGATGTCTGAAGTGGAGAGTCT 1006  
DB 1023 CCCATTGTCCCCCTGTCTATAGGCTTATAGAGTATCTGATGTCTGAAGTGGAGAGTCT 1082  
QY 1007 GTTCTTGAACCTGCTCTTGAATCTGTCAAAAGTTTCAATCCAGGGCCACAGCTTACAGCC 1066  
DB 1083 GTTCTTGAACCTGCTCTTGAATCTGTCAAAAGTTTCAATCCAGGGCCACAGCTTACAGCC 1142  
QY 1067 ACTCCAAATTAAGATGTCATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1126  
DB 1143 ACTCCAAATTAAGATGTCATTAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1202  
QY 1127 CTCTGTGATGCAATTCATATTGGGATTCGCAATGGGCGACATTAATAATCCAAATCC 1186  
DB 1203 CTCTGTGATGCAATTCATATTGGGATTCGCAATGGGCGACATTAATAATCCAAATCC 1262  
QY 1187 CACTTGAACCACTGAAGAAAGAGATTTGACTCAGATGCTGTCAACACCATATGAA 1246  
DB 1263 CACTTGAACCACTGAAGAAAGAGATTTGACTCAGATGCTGTCAACACCATATGAA 1322  
QY 1247 AGTCAGAGTGTTCCTCCAGCTATTAACAAAAGTGAAGGAGGATCCCGAGGCGAG 1306  
DB 1323 AGTCAGAGTGTTCCTCCAGCTATTAACAAAAGTGAAGGAGGATCCCGAGGCGAG 1382  
QY 1307 AATGATCAAGAGCTGAAGTATCAAGTATTAAGAGACATGTCCAGTGGGCTTTTGAAGAGTG 1366  
DB 1383 AATGATCAAGAGCTGAAGTATCAAGTATTAAGAGACATGTCCAGTGGGCTTTTGAAGAGTG 1442  
QY 1367 GTGGGATCCAGTTCAGAGAGAGGAGTATGTTGTCTCCAGTGGGCGAAGAGTGC 1426  
DB 1443 GTGGGATCCAGTTCAGAGAGAGGAGTATGTTGTCTCCAGTGGGCGAAGAGTGC 1502  
QY 1427 TATGCGGAATTCCTGCTATGACAGAAAGTCCACCATTTCTTGTGATGATGCTTTAA 1486  
DB 1503 TATGCGGAATTCCTGCTATGACAGAAAGTCCACCATTTCTTGTGATGATGCTTTAA 1562  
QY 1487 AGTCTACGTTCTCTAATATAGAAACAGCAGTCTTGTGACGCTTGTGTGGCTGATG 1546  
DB 1563 AGTCTACGTTCTCTAATATAGAAACAGCAGTCTTGTGACGCTTGTGTGGCTGATG 1622  
QY 1547 GTCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1606  
DB 1623 GTCTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1682  
QY 1607 TTAATTAAGACACACAGATTCACATTTTATACATGAGAGATCTTGTGTGATGATG 1666  
DB 1683 TTAATTAAGACACACAGATTCACATTTTATACATGAGAGATCTTGTGTGATGATG 1742  
QY 1667 CCAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1726  
DB 1743 CCAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1801  
QY 1727 CTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATGATGATGATG 1786  
DB 1802 CTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATGATGATGATG 1861  
QY 1787 TAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1846  
DB 1862 TAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1921  
QY 1847 TAGTCTCGGAATTTACAGAGAGAGGAAATCAACTGAGAGAACTGTGACATGAGAG 1906  
DB 1922 TAGTCTCGGAATTTACAGAGAGAGGAAATCAACTGAGAGAACTGTGACATGAGAG 1981  
QY 1907 TTGAAGACCAAGACTTTGAAATTTGCAAGCTGCTCATGTGTGATTAATATCATGCTG 1966  
DB 1982 TTGAAGACCAAGACTTTGAAATTTGCAAGCTGCTCATGTGTGATTAATATCATGCTG 2041



QY	1723	TTATCTAAATTTCCAGATGCTTTTGTAGATGACTGAAGATTTTGTAGCCACATATTTGGGA	1782
Db	1804	TTATCTAAATTTCCAGATGCTTTTGTAGATGACTGAAGATTTTGTAGCCACATATTTGGGA	1863
QY	1783	GTTCTAGATTTTGAGTGAATGCGAGGAAGGGCCATCTCATTTGAGATGATTTAAGTGAAC	1842
Db	1864	GTTCTAGATTTTGAGTGAATGCGAGGAAGGGCCATCTCATTTGAGATGATTTAAGTGAAC	1923
QY	1843	AAACTAGTTCTCGGAATTTCTACAGAGAAGGAGGAATCAGACTGAGGAAGCTGTGACATA	1902
Db	1924	AAACTAGTTCTCGGAATTTCTACAGAGAAGGAGGAATCAGACTGAGGAAGCTGTGACATA	1983
QY	1903	GGACTTGAAGCCAAAGACTTTGAAATTTTGGAGCTGCTCATATGTGTGATTAATATTCAC	1962
Db	1984	GGACTTGAAGCCAAAGACTTTGAAATTTTGGAGCTGCTCATATGTGTGATTAATATTCAC	2043
QY	1963	GCTGCTTTCTATTTGAGTTACAAATCTATATTTTATTTGAGATTTAAATTAAGAAAAAT	2022
Db	2044	GCTGCTTTCTATTTGAGTTACAAATCTATATTTTATTTGAGATTTAAATTAAGAAAAAT	2103
QY	2023	TTACAGAGAAAAAATTT	2041
Db	2104	TTAAAAAATTT	2122
RESULT 6			
US-10-133-013-165			
; Sequence 165, Application US/10133013			
; GENERAL INFORMATION:			
; APPLICANT: Astronomoff, Anna			
; APPLICANT: Bandman, Olga			
; APPLICANT: Cocks, Benjamin G.			
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE			
; FILE REFERENCE: PA-0049 US			
; CURRENT APPLICATION NUMBER: US/10/133.013			
; CURRENT FILING DATE: 2002-04-25			
; PRIOR APPLICATION NUMBER: 60/287,067			
; PRIOR FILING DATE: 2001-04-27			
; NUMBER OF SEQ ID NOS: 271			
; SOFTWARE: PERL Program			
; SEQ ID NO 165			
; LENGTH: 2108			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc-feature			
; OTHER INFORMATION: Incyte ID No: 977391.15			
US-10-133-013-165			
Query Match			
Best Local Similarity 96.1%; Pred. No. 0;			
Matches 2025; Conservative 0; Mismatches 4; Indels 79; Gaps			
QY	1	CTGCCATTAAGATGGCGTCCGTGGCGGCTGCAAGACAGTTCCTGTGGCAGTGGGCTCAG	60
Db	1	CTGCCATTAAGATGGCGTCCGTGGCGGCTGCAAGACAGTTCCTGTGGCAGTGGGCTCAG	60
QY	61	GGCGCTGCAAGGAGACCTCACTCTGTGTAGTATTTCTGGGGGCAAGGGCACCGGAAATC	120
Db	61	GGCGCTGCAAGGAGACCTCACTCTGTGTAGTATTTCTGGGGGCAAGGGCACCGGAAATC	120
QY	121	CACGCTGCAGTTGCGAGCTAGGCGCAGCGGCTGGCGGTGAGATCGTCACGCTGACTCCAT	180
Db	121	CACGCTGCAGTTGCGAGCTAGGCGCAGCGGCTGGCGGTGAGATCGTCACGCTGACTCCAT	180
QY	181	GCAGGCTATGGAAGGCTTAGACATCATCAACCAAGGTTTCTGCCCAAGACAGAGAT	240
Db	181	GCAGGCTATGGAAGGCTTAGACATCATCAACCAAGGTTTCTGCCCAAGACAGAGAT	240
QY	241	CTGCCGGACACATGATCAGTCTTGTGTGATTCCTCTTGTGACCAATTACAGAGTGTGGA	300
Db	241	CTGCCGGACACATGATCAGTCTTGTGTGATTCCTCTTGTGACCAATTACAGAGTGTGGA	300

Db 1445 GGTGGTGGGATCCAGTTCAGAGAGGAGGCTATGTTCTCCCATCTCGGCAAGGA 1504  
1423 GTGCTATGCGGAATTCCTGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGT 1482  
1505 GTGCTATGCGGAATTCCTGCATAGCAGAAAGCTCCACCATTTCTTTGATGTGT 1564  
1483 TTAAGTCTCAGCTTCTCTATATAGAAACAGAGCTGTGACGCTCTGTGTGCTG 1542  
1565 TTAAGTCTCAGCTTCTCTATATAGAAACAGAGCTGTGACGCTCTGTGTGCTG 1624  
1543 ATGAGTCTGGAATGATGTTGTTGAGAAAGCATTTTCTTTTGAACCTTAAGT 1602  
1525 ATGAGTCTGGAATGATGTTGTTGAGAAAGCATTTTCTTTTGAACCTTAAGT 1664  
1603 TCTATTTTAAAGCAGACAGATTCACATTTTATACATGAGAGATCTTCTGTG 1662  
1685 TCTATTTTAAAGCAGACAGATTCACATTTTATACATGAGAGATCTTCTGTG 1744  
1663 AATACCGAGATTCAGTCCCTTTAAAGAACTTTATGCTCCCTGACTGCTGCTAAA 1722  
1745 AATACCGAGATTCAGTCCCTTTAAAGAACTTTATGCTCCCTGACTGCTGCTAAA 1803  
1723 TTAATCTATTTCCAGATGCTTTGTAGATGAGTATGAGGACATATTTGGGA 1782  
1804 TTAATCTATTTCCAGATGCTTTGTAGATGAGTATGAGGACATATTTGGGA 1863  
1783 GTTCTAGATTTGAGTGAATGGCAGAAAGGCGCATCTCCATGAGATTAAGTGAAC 1842  
1864 GTTCTAGATTTGAGTGAATGGCAGAAAGGCGCATCTCCATGAGATTAAGTGAAC 1923  
1843 AACTAGTCTCGGAATTCACAGAAAGGAGGAGTATGAGTGAAGCTGTGACATA 1902  
1924 AACTAGTCTCGGAATTCACAGAAAGGAGGAGTATGAGTGAAGCTGTGACATA 1983  
1903 GAGCTGAAGACCAAGACTTTGAATTTGAGAGCTGCTATGCTGATTAATCACT 1962  
1984 GAGCTGAAGACCAAGACTTTGAATTTGAGAGCTGCTATGCTGATTAATCACT 2043  
1963 GCTGCTTTTATGAGTGAATCAATCTATTTTATGAGTTAAATTAAGAAAAAT 2022  
2044 GCTGCTTTTATGAGTGAATCAATCTATTTTATGAGTTAAATTAAGAAAAAT 2103  
1963 GCTGCTTTTATGAGTGAATCAATCTATTTTATGAGTTAAATTAAGAAAAAT 2022  
2044 GCTGCTTTTATGAGTGAATCAATCTATTTTATGAGTTAAATTAAGAAAAAT 2103  
QY 2023 TTACAGAGAAAAAATAA 2041  
Db 2104 TTAAGAAAAAATAAATAA 2122

RESULT 5  
US-10-097-340-150  
; Sequence 150, Application US/10097340  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNANVARAPU  
; APPLICANT: Sebastian HOERSCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANDT  
; APPLICANT: Xunel ZHAO  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14

PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/325,102  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/323,580  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ. ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ. ID NO 150  
; LENGTH: 2129  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-097-340-150

Query Match 94.5%; Score 1929.4; DB 39; Length 2129;  
Best Local Similarity 95.8%; Pred. No. 0;  
Matches 2029; Conservative 5; Mismatches 6; Indels 79; Gaps 2;

QY 1 CTGCGATTAAGTGGCGTCCGTCGGGCGTCGACAGACATTCCTGTTGGGACATGGGCTCAG 60  
Db 5 CTGCGATTAAGTGGCGTCCGTCGGGCGTCGACAGACATTCCTGTTGGGACATGGGCTCAG 64  
QY 61 GGGCGTCGACAGACATTCCTGTTGGGCGTCGACAGACATTCCTGTTGGGACATGGGCTCAG 120  
Db 65 GGGCGTCGACAGACATTCCTGTTGGGCGTCGACAGACATTCCTGTTGGGACATGGGCTCAG 124  
QY 121 CACGCTGGCGTTGACGATGAGCGCAGCGCTCGCGGCTGAGATCGTACGCTGACTCAT 180  
Db 125 CACGCTGGCGTTGACGATGAGCGCAGCGCTCGCGGCTGAGATCGTACGCTGACTCAT 184  
QY 181 GCAGCTATTAAGGCGCTAGACATCATCACAACAAGTTCTGCGCAAGAGCAGAGAT 240  
Db 185 GCAGCTATTAAGGCGCTAGACATCATCACAACAAGTTCTGCGCAAGAGCAGAGAT 244  
QY 241 CTGCGGCGCACATGATGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGTGGA 300  
Db 245 CTGCGGCGCACATGATGAGCTTTGTGATCTCTTGTGACCAATTAACAGTGTGGA 304  
QY 301 CTTCAAGAAATAGACGACGCTCTGATGAGATATTTTGGCCGAGCAAAATCCAT 360  
Db 305 CTTCAAGAAATAGACGACGCTCTGATGAGATATTTTGGCCGAGCAAAATCCAT 364  
QY 361 TGTGTGGAGAACCAATTTATACATGAATCTGCTGAGAAAGTTCTTGTCAATAC 420  
Db 365 TGTGTGGAGAACCAATTTATACATGAATCTGCTGAGAAAGTTCTTGTCAATAC 424  
QY 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGAAGCTTGAAGAA 480  
Db 425 CAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGAAGCTTGAAGAA 484  
QY 481 GGATGCTTTGACTTCAACAAAGCTTAAAGCAGGTGAGCAAAATGGCTCCCACT 540  
Db 485 GGATGCTTTGACTTCAACAAAGCTTAAAGCAGGTGAGCAAAATGGCTCCCACT 544  
QY 541 GCATCCATGACCAAGCAAGTGGCCAGAGCTTGCAAGTTTGAAGAAAGGAGAT 600  
Db 545 GCATCCATGACCAAGCAAGTGGCCAGAGCTTGCAAGTTTGAAGAAAGGAGAT 604  
QY 601 CTTCTATGATGATTTCTCATGCTCAACATCGGAAAGAGTGGTGTCCCTTGGAGG 660  
Db 605 CTTCTATGATGATTTCTCATGCTCAACATCGGAAAGAGTGGTGTCCCTTGGAGG 664  
QY 661 TCCTCTGAAGTCTCTAACCCTTGATCTTGGCTTCACTGCTGACAGGACATTTCTAGA 720  
Db 665 TCCTCTGAAGTCTCTAACCCTTGATCTTGGCTTCACTGCTGACAGGACATTTCTAGA 724

QY	1903	GGACTTGAAGCCAAAGACTTTGAAATTTGCGACCTGCTCAGCTGTGAGTATTATCACT	1962
Db	1984	GGACTTGAAGACCAAGACTTTGAAATTTGCGACCTGCTCAGCTGTGAGTATTATCACT	2043
QY	1963	GCTGCTTTCATATGATGATTAACAATCTATATTTTATGAGTTAAATAAGAAAAAT	2022
Db	2044	GCTGCTTTCATATGATGATTAACAATCTATATTTTATGAGTTAAATAAGAAAAAT	2103
QY	2023	TTACACGAAAAAAAAAAAAA	2041
Db	2104	TTAAAAAAAAAAAAAAAAAA	2122

```

RESULT 4
PCT-0502-07826-150
; Sequence 150, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
; OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 150
; LENGTH: 2129
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-0502-07826-150

```

Query Match	94.5%	Score 1929.4	DB 1	Length 2129
Best Local Similarity	95.8%	Pred. No. 0		
Matches 2029	Conservative	5	Mismatches	6
			Indels	79
			Gaps	2
Qy	1	CTGCATTAAGATGGGCGTCCGTGGCGAGCTTCACAGACAGTTCCTGTTGGGCGAGTGGGCTCAG	60	
Db	5	CTGCATTAAGATGGGCGTCCGTGGCGCTTCACAGACAGTTCCTGTTGGGCGAGTGGGCTCAG	64	
Qy	61	GGGCGCTGCACGCGACCTTACCTCTTGTGTAGTATTCTCGGGGCGACGGGCGACAATC	120	
Db	65	GGGCGCTGCACGCGACCTTACCTCTTGTAGTATTCTCGGGGCGACGGGCGACSSGMRATC	124	
Qy	121	CACGCTGGCGGTTGCACGCTAGGCGCAGGGGCGTGGCGGTGAGATCGTACGCGTACTCCAT	180	
Db	125	CACGCTGGCGGTTGCACGCTAGGCGCAGGGGCTGGCGGTGAGATCGTACGCGTACTCCAT	184	
Qy	181	GCAGGCTCATGGAAGGCTGAGACATATCCACCAACAAGTTCCTGCCCCAAGACAGAGAAAT	240	
Db	185	GCAGGCTCATGGAAGGCTGAGACATATCCACCAACAAGTTCCTGCCCCAAGACAGAGAAAT	244	
Qy	241	CTGCCGGGCGACCAATGATAGCTTTTGGGATCCTCTTGTGACCAATATCACAGTGTGGA	300	
Db	245	CTGCCGGGCGACCAATGATAGCTTTTGGGATCCTCTTGTGACCAATATCACAGTGTGGA	304	
Qy	301	CTTCGAAATTAAGCAACGCTCTGATGATGAATATATTTGGCCGAGCAAAATTCCTAT	360	
Db	305	CTTCGAAATTAAGCAACGCTCTGATGATGAATATATATTTGGCCGAGCAAAATTCCTAT	364	

QY	361	TG	GTGGGAGAGAACCAATTATTAACATGTGAATCTCGCTCGGAAAGTCTTGTCATATAC	420
Db	365	TG	TGTGTGGAGAGAACCAATTATTAACATGTGAATCTCGCTCGGAAAGTCTTGTCATATAC	424
QY	421	CA	AGGCCCGAGAGATGTGGGACATGAGAAAGATGTGACCAAAAGTGTGAGCTTGAAGAAAGA	480
Db	425	CA	AGGCCCGAGAGATGTGGGACATGAGAAAGATGTGACCAAAAGTGTGAGCTTGAAGAAAGA	484
QY	481	GG	ATGTGCTTTGTACTTCACAAAGCCCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGCT	540
Db	485	GG	ATGTGCTTTGTACTTCACAAAGCCCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGCT	544
QY	541	GC	ATTCACATATACAAAGCAAGCAAGTGGCGAGACCTTGCAAGTTTGTGAAGAAACAGCAAT	600
Db	545	GC	ATTCACATATACAAAGCAAGCAAGTGGCGAGACCTTGCAAGTTTGTGAAGAAACAGCAAT	604
QY	601	CT	CTCATATGTAATTTCTCATGCTGCAACATACGAGAAGAGTGGTGGTCCCTGTGAGG	660
Db	605	CT	CTCATATGTAATTTCTCATGCTGCAACATACGAGAAGAGTGGTGGTCCCTGTGAGG	664
QY	661	TC	CTCTGGAAGTTCTCTTAACCCCTTGCAATCCTTGGCTTCATGCTGACACAGGACGTTCTAGA	720
Db	665	TC	CTCTGGAAGTTCTCTTAACCCCTTGCAATCCTTGGCTTCATGCTGACACAGGACGTTCTAGA	724
QY	721	TG	AGGCGCTTGGATTAAGAGGGTGGATGACATGCTGCTGCTGGGCTCTTGGAGAACTAAG	780
Db	725	TG	AGGCGCTTGGATTAAGAGGGTGGATGACATGCTGCTGCTGGGCTCTTGGAGAACTAAG	784
QY	781	AG	ATTTTACAGACGCGATATATACGAAGAATGTTTGGGAAAAATAGCAGACATATACACA	840
Db	785	AG	ATTTTACAGACGCGATATATACGAAGAATGTTTGGGAAAAATAGCAGACATATACACA	844
QY	841	TG	GTATATCTCCAAATCAATTTGGCTTCAAGAAATTTACAGATACCTGATACCTAGAGGAAA	900
Db	845	TG	GTATATCTCCAAATCAATTTGGCTTCAAGAAATTTACAGATACCTGATACCTAGAGGAAA	904
QY	901	AT	GCACTGGAAGCTAGTAACACGCTTCTAAGAAAG-----	938
Db	905	AT	GCACTGGAAGCTAGTAACACGCTTCTAAGAAAGGATTTGAGGCTGTGAACCAAGT	964
QY	939	-----	-----GAGC	942
Db	965	AA	CTAAGAGATATGCCCGGAAACAAAACGATGGGTTAAAAACCGTTTGTGAGCAGACC	1024
QY	943	TG	GTGTCATTTGCCCCCTGTCTATGTGCTTAGAGGTATGATGTCTCGAATGGGAGGA	1002
Db	1025	TG	GTGTCATTTGCCCCCTGTCTATGTGCTTAGAGGTATGATGTCTCGAATGGGAGGA	1008
QY	1003	GT	CTGTGTTCTTAACCTGCTCTTGAATCTGTGCAAAAGTTTCATCCAGGGCCACACGCTTAC	1062
Db	1085	GT	CTGTGTTCTTAACCTGCTCTTGAATCTGTGCAAAAGTTTCATCCAGGGCCACACGCTTAC	1144
QY	1063	AG	CCATCAATTAAGATGGCATATGAAAGGTGAGAAACAGAAAGTTATCACCTGTG	1122
Db	1145	AG	CCATCAATTAAGATGGCATATGAAAGGTGAGAAACAGAAAGTTATCACCTGTG	1204
QY	1123	TG	ACCTCTGTATGTGAATCATATTTGGGATTCGCGAATGGGCGACGACATATAATCCAA	1182
Db	1205	TG	ACCTCTGTATGTGAATCATATTTGGGATTCGCGAATGGGCGACGACATATAATCCAA	1264
QY	1183	AT	CCCATCTTGAACCACTGGAAGAAAGAGAGATTGGACTCAGATGCTGTACACCAT	1242
Db	1265	AT	CCCATCTTGAACCACTGGAAGAAAGAGAGATTGGACTCAGATGCTGTACACCAT	1324
QY	1243	AG	AAAGTCAAGATGTTTCCCGACGCTTAACAAAGAAACCTAAAGGGAAGGATGCCCAAG	1302
Db	1325	AG	AAAGTCAAGATGTTTCCCGACGCTTAACAAAGAAACCTAAAGGGAAGGATGCCCAAG	1384
QY	1303	GC	AGATATCAAGAAGCTGAATTCACACGTTTAAAGAGACATGTCCAGTGGCTTTGGAAA	1362
Db	1385	GC	AGATATCAAGAAGCTGAATTCACACGTTTAAAGAGACATGTCCAGTGGCTTTGGAAA	1444
QY	1363	GG	TGTGTGGGATTCAGATTCAAGAGAGGGGTATGTTTGTCTCCAGTCTGGGCAAGGA	1422

```

; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 152
; LENGTH: 2129
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-152

Query Match      94.6%; Score 1931; DB 39; Length 2129;
Best Local Similarity 95.9%; Pred: No 0;
Matches 2033; Conservative 1; Mismatches 6; Indels 79; Gaps 2;

QY      1 CTGCCATAGATGGCGTCCGTGGCGGTGACAGAGCAGTTCTGTGGGAGTGGGCTCAG 60
        5 CTGCCATAGATGGCGTCCGTGGCGGTGACAGAGCAGTTCTGTGGGAGTGGGCTCAG 64
        61 GGGCCCTGCAAGCAGCCTTACTCTGTGTAGTATTTCTGGGGCCACGGGCAACGCAATC 120
        65 GGGCCCTGCAAGCAGCCTTACTCTGTGTAGTATTTCTGGGGCCACGGGCAACGCAATC 124
QY      121 CAGCGTGGCTGACGTAGCGCCAGCGGCTGGGGGTGAGATGCTGAGCGCTGACTCCAT 180
        125 CAGCGTGGCTGACGTAGCGCCAGCGGCTGGGGGTGAGATGCTGAGCGCTGACTCCAT 184
QY      181 GCAGGCTATGAGGCGCTAGACATCATCACCAACAGGTTTCTGCCAGAGAGAGAT 240
        185 GCAGGCTATGAGGCGCTAGACATCATCACCAACAGGTTTCTGCCAGAGAGAGAT 244
QY      241 CTGCGCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGA 300
        245 CTGCGCGGACACCATGATGATGCTTTGTGATCTCTTGTGACCAATTACACAGTGTGA 304
QY      301 CTTCAGAAATAGAGCAACTGCTGATTTGAAGATATTTGGCCGAGACAAAATTCCTAT 360
        305 CTTCAGAAATAGAGCAACTGCTGATTTGAAGATATTTGGCCGAGACAAAATTCCTAT 364
QY      361 TGTGTGGAGAGAACCAATTATTACATTGAATCTGCTGTGAAAATTTCTGTCAATAC 420
        365 TGTGTGGAGAGAACCAATTATTACATTGAATCTGCTGTGAAAATTTCTGTCAATAC 424
QY      421 CAGGCCCGAGAGAGGCGCTGAGAAAGTGAACCGGAAAGTGGAGCTTGAAGAAGA 480
        425 CAGGCCCGAGAGAGGCGCTGAGAAAGTGAACCGGAAAGTGGAGCTTGAAGAAGA 484
QY      481 GGATGGTCTGTACTTACAAAGCGCTAAGCCAGGTGGACCGAAGATGGCTGCCAAGCT 540
        485 GGATGGTCTGTACTTACAAAGCGCTAAGCCAGGTGGACCGAAGATGGCTGCCAAGCT 544
QY      541 GCATCCACATGACAAACGCAAGTGGCGAGAGCTTGCAGATTTTGAAGAAACAGGAAT 600
        545 GCATCCACATGACAAACGCAAGTGGCGAGAGCTTGCAGATTTTGAAGAAACAGGAAT 604
QY      601 CTTCATAGGAAATTTCTCATGTCACATAGGAAAGAGTGGTGGTCCCTTGGAGG 660
        605 CTTCATAGGAAATTTCTCATGTCACATAGGAAAGAGTGGTGGTCCCTTGGAGG 664
QY      661 TCCCTGAAATTTCTTAACCTTGCATCTTGGCTCATGTCAGCAGGAGAGTTCAGA 720
        665 TCCCTGAAATTTCTTAACCTTGCATCTTGGCTCATGTCAGCAGGAGAGTTCAGA 724
QY      721 TGAGCGCTGGATPAAGAGGTGATGATGCTGCTGGGGCTTGGAGGAAGTAAAG 780
        725 TGAGCGCTGGATPAAGAGGTGATGATGCTGCTGGGGCTTGGAGGAAGTAAAG 784
QY      781 AGATTTTCAAGAGCGTATATCAAGAAAGATGTTGGGAAATATAGCCAGACATATACA 840
        785 AGATTTTCAAGAGCGTATATCAAGAAAGATGTTGGGAAATATAGCCAGACATATACA 844
QY      841 TGGTATCTTCCATCAATTTGGCTTCAAGAAATTTCAAGAGTACCTGATCAGAGGAAA 900
        845 TGGTATCTTCCATCAATTTGGCTTCAAGAAATTTCAAGAGTACCTGATCAGAGGAAA 904
```

```

QY      901 ATGCACACTGGAGACTAGTAACCAAGCTTCTAAGAAAG----- 938
        905 ATGCACACTGGAGACTAGTAACCAAGCTTCTAAGAAAGGATTTAGGCTGTGAACAAAGT 964
QY      939 -----GACC 942
        965 AACTAAGAGATATGCCCGGAAACAAACCGATGGGTTAAACCGTTTTTGAAGCACACC 1024
QY      943 TGGTCCATGTCGCCCTGCTATGAGCTTAGAGGATATGATGTCGAAAGTGGGAGA 1002
        1025 TGGTCCATGTCGCCCTGCTATGAGCTTAGAGGATATGATGTCGAAAGTGGGAGA 1084
QY      1003 GTCTGTTCTTGAACCTGCTTTGAATCGTCAAAAGTTTATCCAGGCGCAAGCCTAC 1062
        1085 GTCTGTTCTTGAACCTGCTTTGAATCGTCAAAAGTTTATCCAGGCGCAAGCCTAC 1144
QY      1063 AGCCACTCCAAATAAGATGCCATACATGAACTGAGAACACAGAAAGTTATCACTGTG 1122
        1145 AGCCACTCCAAATAAGATGCCATACATGAACTGAGAACACAGAAAGTTATCACTGTG 1204
QY      1123 TGAACCTGTGATGCAATCATATTTGGGGATCGGGAATGGGCGACATAAATCCAA 1182
        1205 TGAACCTGTGATGCAATCATATTTGGGGATCGGGAATGGGCGACATAAATCCAA 1264
QY      1183 ATCCACTTGAACCAACTGAAAGAAAGAAAGATTTGACTCAGATGCTGTCAACACAT 1242
        1265 ATCCACTTGAACCAACTGAAAGAAAGAAAGATTTGACTCAGATGCTGTCAACACAT 1324
QY      1243 AGAAAGTCAAGATGTTTCCCGAGACTATACAAAGAACTTAAGGAGAGGATCCCAAG 1302
        1325 AGAAAGTCAAGATGTTTCCCGAGACTATACAAAGAACTTAAGGAGAGGATCCCAAG 1384
QY      1303 GCAGAAATGATCAAGAGCTGAATGCAGCTTTAAGACATGTCATGAGCTTTGGAAA 1362
        1385 GCAGAAATGATCAAGAGCTGAATGCAGCTTTAAGACATGTCATGAGCTTTGGAAA 1444
QY      1363 GGTGTGGGGATCCAGTTTCAAGAGGAGGATATGTTGTCTCCCACTGGGCAAGA 1422
        1445 GGTGTGGGGATCCAGTTTCAAGAGGAGGATATGTTGTCTCCCACTGGGCAAGA 1504
QY      1423 GTGCTATGCGGAATTTCTCTCATAGCAGAAAGCTCCACCATTTTGTGATGTGTT 1482
        1505 GTGCTATGCGGAATTTCTCTCATAGCAGAAAGCTCCACCATTTTGTGATGTGTT 1564
QY      1483 TTAAGTCTCAGCTTCTTAATATAGAAACAGAGCTTGTGAGCTGCTGTGTGCTG 1542
        1565 TTAAGTCTCAGCTTCTTAATATAGAAACAGAGCTTGTGAGCTGCTGTGTGCTG 1624
QY      1543 ATGTGTCTGGAATGATGTAGTTCAGAAAGCAATTTTTTTTCTTGAACCTTAAAGT 1602
        1625 ATGTGTCTGGAATGATGTAGTTCAGAAAGCAATTTTTTTTCTTGAACCTTAAAGT 1684
QY      1603 TCTATTTTAAAGCAGCAAGATTCACATTTTATACATGAGAGATCTTTGTGTG 1662
        1685 TCTATTTTAAAGCAGCAAGATTCACATTTTATACATGAGAGATCTTTGTGTG 1744
QY      1663 AATACAGAGATTAAGTCACTCCCTTTAAAGAGTTTATGCTCCCTGACTGCTGCTAAA 1722
        1745 AATACAGAGATTAAGTCACTCCCTTTAAAGAGTTTATGCTCCCTGACTGCTGCTAAA 1803
QY      1723 TTAATCAATTTCCAGATGCTTTGTAGATAGTGAAGTATTTGTGACCAATATTGGA 1782
        1804 TTAATCAATTTCCAGATGCTTTGTAGATAGTGAAGTATTTGTGACCAATATTGGA 1863
QY      1783 GTTCTGATTTGAGTGAATGGAGAAAGGCCATCTCCATTTAGATGATTAAGTGAACC 1842
        1864 GTTCTGATTTGAGTGAATGGAGAAAGGCCATCTCCATTTAGATGATTAAGTGAACC 1923
QY      1843 AAACATGTTCTCGGAATTTCAACAGAGAGAGGAAATCAACTGAGAGAGCTGTGACATA 1902
        1924 AAACATGTTCTCGGAATTTCAACAGAGAGAGGAAATCAACTGAGAGAGCTGTGACATA 1983
```



QY 541 GCATCCACATGACAAAGCGAAAGTGGCAGAGCTTGCAAGTTTGAACAAAGCAAT 600  
 |||||  
 Db 545 GCATCCACATGACAAAGCGAAAGTGGCAGAGCTTGCAAGTTTGAACAAAGCAAT 604  
 QY 601 CTCTCAATGAAATTTCTCCATGCTCAACATACGGAAGAGTGGTCCCTTGGAGG 660  
 |||||  
 Db 605 CTCTCAATGAAATTTCTCCATGCTCAACATACGGAAGAGTGGTCCCTTGGAGG 664  
 QY 661 TCTCTGAAAGTCTCTCAACCTTGCAATCTTGGCTTCAATGCTGACAGGCACTTCA 720  
 |||||  
 Db 665 TCTCTGAAAGTCTCTCAACCTTGCAATCTTGGCTTCAATGCTGACAGGCACTTCA 724  
 QY 721 TGAGCGCTTGGAATGAAGGGGTGATGACATGCTGCTGGGCTCTTGAGGAACCTAG 780  
 |||||  
 Db 725 TGAGCGCTTGGAATGAAGGGGTGATGACATGCTGCTGGGCTCTTGAGGAACCTAG 784  
 QY 781 AGATTTTACAGAGCGCTATATACAGAAAGATGTTGGAAGAAATAGCCAGAGCTATCA 840  
 |||||  
 Db 785 AGATTTTACAGAGCGCTATATACAGAAAGATGTTGGAAGAAATAGCCAGAGCTATCA 844  
 QY 841 TGGTATCTTCCAAATCAATGGCTTCAAGAAATTTACAGAGTACCTGATCACTGAGGAAA 900  
 |||||  
 Db 845 TGGTATCTTCCAAATCAATGGCTTCAAGAAATTTACAGAGTACCTGATCACTGAGGAAA 904  
 QY 901 ATGCACACTGGAGACTAGTAACCAAGCTTCTAAGAAAG----- 938  
 |||||  
 Db 905 ATGCACACTGGAGACTAGTAACCAAGCTTCTAAGAAAGATATTGAGGCTCTGAACAAGT 964  
 QY 939 -----GACC 942  
 |||||  
 Db 965 AACTAAGAGATATGCCCGGAAACAAACGATGGTTAAACCGTTTTTTAGACAGACC 1024  
 QY 943 TGGTCCCATTTGCCCTGCTCTATAGGCTTAGAGGTTCTGATGCTCGAAGTGGAGGA 1002  
 |||||  
 Db 1025 TGGTCCCATTTGCCCTGCTCTATAGGCTTAGAGGTTCTGATGCTCGAAGTGGAGGA 1084  
 QY 1003 GTCCTGTTGTAACGCTGCTTGAATCGTGAAGTTTCATCCAGGCGCAACAGCTTAC 1062  
 |||||  
 Db 1085 GTCCTGTTGTAACGCTGCTTGAATCGTGAAGTTTCATCCAGGCGCAACAGCTTAC 1144  
 QY 1063 AGCCACTCCAAATTAAGATGCTATACATGAGTGAACAAAGAAATTTATCAGCTGTG 1122  
 |||||  
 Db 1145 AGCCACTCCAAATTAAGATGCTATACATGAGTGAACAAAGAAATTTATCAGCTGTG 1204  
 QY 1123 TGAAGCTGCTGATGATCAATCATCTTGGGATCGCGAATGGGCGACATATAATTCAA 1182  
 |||||  
 Db 1205 TGAAGCTGCTGATGATCAATCATCTTGGGATCGCGAATGGGCGACATATAATTCAA 1264  
 QY 1183 ATCCCACTTGAACCAACTGAAGAAAGAAAGAAATTTGAGTCAAGTCTTCAACACAT 1242  
 |||||  
 Db 1265 ATCCCACTTGAACCAACTGAAGAAAGAAAGAAATTTGAGTCAAGTCTTCAACACAT 1324  
 QY 1243 AGAAAGTCAAGAGTGTTCCTCCAGACTATTAACAAGAAACCTTAAGGAGGATCCCGAG 1302  
 |||||  
 Db 1325 AGAAAGTCAAGAGTGTTCCTCCAGACTATTAACAAGAAACCTTAAGGAGGATCCCGAG 1384  
 QY 1303 GCAGAAATGATCAAGAGCTGAATGACAGCTTTAAGAGACATGTCAGTGGCTTTGAAA 1362  
 |||||  
 Db 1385 GCAGAAATGATCAAGAGCTGAATGACAGCTTTAAGAGACATGTCAGTGGCTTTGAAA 1444  
 QY 1363 GGTGGTGGGATGCCAATTTGAGGAGGAGGTATGTTTGTCTCCAGTCTGGGCAAGGA 1422  
 |||||  
 Db 1445 GGTGGTGGGATGCCAATTTGAGGAGGAGGTATGTTTGTCTCCAGTCTGGGCAAGGA 1504  
 QY 1423 GTGCTATGGGAATTTCTGATAGCAGAAAAGCTCCCAACATTTTCTTTGATGTTG 1482  
 |||||  
 Db 1505 GTGCTATGGGAATTTCTGATAGCAGAAAAGCTCCCAACATTTTCTTTGATGTTG 1564  
 QY 1483 TTAAGTCTACAGTCTCTATAATAGAAAACAGCAGGCTTGTGAGCTCTTGTGTGCTG 1542  
 |||||  
 Db 1565 TTAAGTCTACAGTCTCTATAATAGAAAACAGCAGGCTTGTGAGCTCTTGTGTGCTG 1624  
 QY 1543 ATGTGCTGTGAAATGATAGTTCAAGAAAAGCATTTTTTTTTTTGAACCTTAAAGGT 1602

Db 1625 ATGTGCTGTGAAATGATAGTTCAAGAAAAGCATTTTTTTTTTTGAACCTTAAAGGT 1684  
 |||||  
 QY 1603 TCTATATTTAAAGCAGACAGATTCACATTTTATACATGAGGATCTTCTTGTGTG 1662  
 |||||  
 Db 1685 TCTATATTTAAAGCAGACAGATTCACATTTTATACATGAGGATCTTCTTGTGTG 1744  
 QY 1663 AATACAGAGTTCAGTGCATCCCTTTAAAGAAATTTATATGTCCTGACTGTGGTAAA 1722  
 |||||  
 Db 1745 AATACAGAGTTCAGTGCATCCCTTT-AAAGAAATTTATGTCCTGACTGTGGTAAA 1803  
 QY 1723 TTAATCAATTTCCAGATGCTTTTGTGATGATGATGATTTGAGGCCCATATTGGGA 1782  
 |||||  
 Db 1804 TTAATCAATTTCCAGATGCTTTTGTGATGATGATGATTTGAGGCCCATATTGGGA 1863  
 QY 1783 GTTCTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1842  
 |||||  
 Db 1864 GTTCTAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1923  
 QY 1843 AAACCTAGTTTCGGAATTTCTACAGAGAGAGGGAATCAGACTGAGAGAGCTGACATA 1902  
 |||||  
 Db 1924 AAACCTAGTTTCGGAATTTCTACAGAGAGAGGGAATCAGACTGAGAGAGCTGACATA 1983  
 QY 1903 GGAAGTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTGATGATGATGATGATGAT 1962  
 |||||  
 Db 1984 GGAAGTGAAGACCAAGACTTTGAAATTTGCGAGCTGCTGATGATGATGATGATGATGAT 2043  
 QY 1963 GCTGCTTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2022  
 |||||  
 Db 2044 GCTGCTTTCTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2103  
 QY 2023 TTAACAGAAAAA 2041  
 |||||  
 Db 2104 TTAACAGAAAAA 2122  
 |||||  
 RESULT 3  
 US-10-097-340-152  
 ? Sequence 152, Application US/10097340  
 ? GENERAL INFORMATION:  
 ? APPLICANT: John MONAHAN  
 ? APPLICANT: Manjula GANNAVAPU  
 ? APPLICANT: Sebastian HOERSCHE  
 ? APPLICANT: Snubhang KAMATKAR  
 ? APPLICANT: Steve G. KOVATS  
 ? APPLICANT: Rachel E. MEYERS  
 ? APPLICANT: Michael MORRISSEY  
 ? APPLICANT: Peter OLANDT  
 ? APPLICANT: Ami SEN  
 ? APPLICANT: Peter VEIBY  
 ? APPLICANT: Gordon B. MILLS  
 ? APPLICANT: Robert C. BAST, Jr.  
 ? APPLICANT: Karen LU  
 ? APPLICANT: Rosemarie SCHMANDT  
 ? APPLICANT: Xumei ZHAO  
 ? APPLICANT: Karen GLATT  
 ? TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
 ? FILE REFERENCE: MRI-030  
 ? CURRENT APPLICATION NUMBER: US/10/097,340  
 ? PRIOR FILING DATE: 2002-03-14  
 ? PRIOR APPLICATION NUMBER: 60/276,025  
 ? PRIOR FILING DATE: 2001-03-14  
 ? PRIOR APPLICATION NUMBER: 60/325,149  
 ? PRIOR FILING DATE: 2001-09-26  
 ? PRIOR APPLICATION NUMBER: 60/276,026  
 ? PRIOR FILING DATE: 2001-03-14  
 ? PRIOR APPLICATION NUMBER: 60/324,967  
 ? PRIOR FILING DATE: 2001/09/26  
 ? PRIOR APPLICATION NUMBER: 60/311,732  
 ? PRIOR FILING DATE: 2001-08-10  
 ? PRIOR APPLICATION NUMBER: 60/325,102  
 ? PRIOR FILING DATE: 2001-09-26



QY 1021 TCTTGAATCGTCGCAAGTTTCATCCAGGGCCACAAAGCTACAGCCATCCCAATTAAGAT 1080  
 |||||  
 Db 1021 TCTTGAATCGTCGCAAGTTTCATCCAGGGCCACAAAGCTACAGCCATCCCAATTAAGAT 1080  
 QY 1081 GCCATCATGAGCTGAGACAAAGAAATTTATCACCCTGTGTGACCTCTGTATGGAAT 1140  
 |||||  
 Db 1081 GCCATCATGAGCTGAGACAAAGAAATTTATCACCCTGTGTGACCTCTGTATGGAAT 1140  
 QY 1141 CATCATTTGGGATCGGCAATGGGCGAGCCATATAAATCCCAATCCCACTGAACCACT 1200  
 |||||  
 Db 1141 CATCATTTGGGATCGGCAATGGGCGAGCCATATAAATCCCAATCCCACTGAACCACT 1200  
 QY 1201 GAAGAAAGAAGATTGCACTCAGATGCTGCAACACCATTGAAGAGTGTTC 1260  
 |||||  
 Db 1201 GAAGAAAGAAGATTGCACTCAGATGCTGCAACACCATTGAAGAGTGTTC 1260  
 QY 1261 CCCAGACTATACAAAGAACTTAAGGGAAGGATCCCGGCGAGAAATATCAAGACT 1320  
 |||||  
 Db 1261 CCCAGACTATACAAAGAACTTAAGGGAAGGATCCCGGCGAGAAATATCAAGACT 1320  
 QY 1321 GAAATGAGCGTTTAAAGACATGTCAGTGGCTTGAAGGTGGTGGGATTCAGTT 1380  
 |||||  
 Db 1321 GAAATGAGCGTTTAAAGACATGTCAGTGGCTTGAAGGTGGTGGGATTCAGTT 1380  
 QY 1381 CAGAGGAGGAGGTATGTTGTCTCCAGTCTGGCAAGAGAGTGTATCGGAATTC 1440  
 |||||  
 Db 1381 CAGAGGAGGAGGTATGTTGTCTCCAGTCTGGCAAGAGAGTGTATCGGAATTC 1440  
 QY 1441 TGCATAGCAGAAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCACGTTCTC 1500  
 |||||  
 Db 1441 TGCATAGCAGAAAAGCTCCACCATTTCTTTGATGTGTTTAAAGTCACGTTCTC 1500  
 QY 1501 TATAATGAAGACGAGGCTTGTGACGCTCTGTGTGCTGATGTCTGCAATGATG 1560  
 |||||  
 Db 1501 TATAATGAAGACGAGGCTTGTGACGCTCTGTGTGCTGATGTCTGCAATGATG 1560  
 QY 1561 TAGTTCGAAAGCAATTTTTTTTCTTGAACCTTAAGGTTCTATTATTAAGAGAGC 1620  
 |||||  
 Db 1561 TAGTTCGAAAGCAATTTTTTTTCTTGAACCTTAAGGTTCTATTATTAAGAGAGC 1620  
 QY 1621 ACAGATTCACATTTTATACATGAGGATCTTGTGTGTAATACAGAGATTGACTGC 1680  
 |||||  
 Db 1621 ACAGATTCACATTTTATACATGAGGATCTTGTGTGTAATACAGAGATTGACTGC 1680  
 QY 1681 ATCCCTTAAAGAGTTTATGCTGCTGACTGCTGCTAAATTAATTAATTCACAGATG 1740  
 |||||  
 Db 1681 ATCCCTTAAAGAGTTTATGCTGCTGACTGCTGCTAAATTAATTAATTCACAGATG 1740  
 QY 1741 CTTTGTAGATGACTGAGTATTTGTGAGCCACATATTTGGGAGTTCTAGATTTGAGTGA 1800  
 |||||  
 Db 1741 CTTTGTAGATGACTGAGTATTTGTGAGCCACATATTTGGGAGTTCTAGATTTGAGTGA 1800  
 QY 1801 TGGCAGGAAGGGCCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCTCGAATT 1860  
 |||||  
 Db 1801 TGGCAGGAAGGGCCATCTCCATTTGAGATGATTAAGTGAACCAACTAGTTCTCGAATT 1860  
 QY 1861 CTACAGGAAGGGAATGAGCTGAGAGCTGAGACTGAGACTGAGAGCAAGCAAGA 1920  
 |||||  
 Db 1861 CTACAGGAAGGGAATGAGCTGAGAGCTGAGACTGAGACTGAGAGCAAGCAAGA 1920  
 QY 1921 CTTTGAATTTGGAGCTGCTCATGTTGAGTTATTTATCATCTGCTGTTCTATTAGT 1980  
 |||||  
 Db 1921 CTTTGAATTTGGAGCTGCTCATGTTGAGTTATTTATCATCTGCTGTTCTATTAGT 1980  
 QY 1981 TACAAATCTATTTTATTGAAGTTAAATTAAGAAAAATTTTACAGAAAAAAA 2040  
 |||||  
 Db 1981 TACAAATCTATTTTATTGAAGTTAAATTAAGAAAAATTTTACAGAAAAAAA 2040  
 QY 2041 A 2041  
 Db 2041 A 2041

RESULT 2  
 PCT-US02-07826-152  
 ; Sequence 152, Application PC/TUS0207826  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Millenium Pharmaceuticals, Inc. et al.  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
 ; FILE REFERENCE: MRI-030PC  
 ; CURRENT APPLICATION NUMBER: PCT/US02/07826  
 ; PRIOR FILING DATE: 2002-03-14  
 ; PRIOR APPLICATION NUMBER: 60/276,025  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/325,149  
 ; PRIOR FILING DATE: 2001-09-27  
 ; PRIOR APPLICATION NUMBER: 60/276,026  
 ; PRIOR FILING DATE: 2001-03-14  
 ; PRIOR APPLICATION NUMBER: 60/324,967  
 ; PRIOR FILING DATE: 2001/09/26  
 ; PRIOR APPLICATION NUMBER: 60/311,732  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: 60/325,102  
 ; PRIOR FILING DATE: 2001-09-26  
 ; PRIOR APPLICATION NUMBER: 60/323,580  
 ; NUMBER OF SEQ ID NOS: 363  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 152  
 ; LENGTH: 2129  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 PCT-US02-07826-152  
 Query Match 94.6%; Score 1931; DB 1; Length 2129;  
 Best Local Similarity 95.9%; Pred. No. 0;  
 Matches 2033; Conservative 1; Mismatches 6; Indels 79; Gaps 2;  
 QY 1 CTCGCATTAATGAGCGCTGCTGCGGCTGCACGAGCAGTCTGTGGGCGAGGCTCAG 60  
 |||||  
 Db 5 CTCGCATTAATGAGCGCTGCTGCGGCTGCACGAGCAGTCTGTGGGCGAGGCTCAG 64  
 QY 61 GGGCGTGCAGAGGACCTTACCTCTTGTAGTATCTGCGGCGACGCGGCAATC 120  
 |||||  
 Db 65 GGGCGTGCAGAGGACCTTACCTCTTGTAGTATCTGCGGCGACGCGGCAATC 124  
 QY 121 CAGGCTGGGCTGCAGCTAGCCAGGCGGCTGGGCGTGAATGCTACGCGTGCATC 180  
 |||||  
 Db 125 CAGGCTGGGCTGCAGCTAGCCAGGCGGCTGGGCGTGAATGCTACGCGTGCATC 184  
 QY 181 GCAGGCTATGAAGGCTTAGACATCATCACCAAGGTTTCTGCCAAGAGCAGAGAT 240  
 |||||  
 Db 185 GCAGGCTATGAAGGCTTAGACATCATCACCAAGGTTTCTGCCAAGAGCAGAGAT 244  
 QY 241 CTGCGGCGACCATGATGATCACTTGTGATCTCTTGTGACCAATTACACAGTGTGA 300  
 |||||  
 Db 245 CTGCGGCGACCATGATGATCACTTGTGATCTCTTGTGACCAATTACACAGTGTGA 304  
 QY 301 CTTTCAGAAATGAGCACTGCTGATGTAAGATATTTGGCGGAGACAAATTCCTAT 360  
 |||||  
 Db 305 CTTTCAGAAATGAGCACTGCTGATGTAAGATATTTGGCGGAGACAAATTCCTAT 364  
 QY 361 TGTGTGGGAGGAACCAATTAATGATGATCTGCTCTGGAATTTCTGTGATATAC 420  
 |||||  
 Db 365 TGTGTGGGAGGAACCAATTAATGATGATCTGCTCTGGAATTTCTGTGATATAC 424  
 QY 421 CAAGCCCGAGAGATGGGCACTGAGAAAGTGAATGACCGAAAGGTGAGCTTGAAGA 480  
 |||||  
 Db 425 CAAGCCCGAGAGATGGGCACTGAGAAAGTGAATGACCGAAAGGTGAGCTTGAAGA 484  
 QY 481 GGATGGCTGTACTTACAAAGCGCTAACGCCAGGTGGAGCCAGAAATGGCTGCAAGCT 540  
 |||||  
 Db 485 GGATGGCTGTACTTACAAAGCGCTAACGCCAGGTGGAGCCAGAAATGGCTGCAAGCT 544



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:26:48 ; Search time 6028.18 Seconds

(without alignments)  
8512.654 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041  
Sequence: 1 CTGCCATAGATGATGCTCG.....TTTACAGAAAAA 2041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

```

Pending Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/US06000.COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06001.COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US06002.COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US06003.COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US06004.COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US06005.COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US06006.COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US06007.COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US06008.COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US06009.COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US06010.COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US06011.COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US06012.COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US06013.COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US06014.COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US06015.COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US06016.COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US06017.COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US06018.COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US06019.COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US06020.COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US06021.COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US06022.COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US06023.COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US06024.COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US06025.COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US06026.COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US06027.COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US06028.COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US06029.COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US06030.COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US06031.COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US06032.COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US06033.COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US06034.COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US06035.COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US06036.COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US06037.COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US06038.COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US06039.COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US06040.COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US06041.COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US06042.COMB.seq.*

```

Result No.	Score	Query Match	Length	DB ID	Description
1	2041	100.0	2041	19 US-09-513-151-3	Sequence 152, App
2	1931	94.6	2129	1 PCT-US02-07826-112	Sequence 152, App
3	1931	94.6	2129	39 US-10-097-340-152	Sequence 150, App
4	1929.4	94.5	2129	1 PCT-US02-07826-110	Sequence 150, App
5	1929.4	94.5	2129	39 US-10-097-340-150	Sequence 150, App
6	1923.6	94.2	2108	40 US-10-133-013-165	Sequence 165, App
7	1909	92.8	2109	76 US-60-324-185-33690	Sequence 33690, A
8	1894.6	92.8	2101	61 US-60-172-360-26034	Sequence 26034, A
9	1646	82.6	1858	18 US-09-471-275-28104	Sequence 2814, Ap
10	1556.8	76.3	1749	1 PCT-US02-18947-1692	Sequence 1692, Ap
11	1556.8	76.3	1749	41 US-10-172-118-1692	Sequence 1692, Ap
12	1088.2	53.3	1183	29 US-09-757-028-874	Sequence 874, App
13	1088.2	53.3	1183	42 US-10-222-911-874	Sequence 874, App
14	1065.6	52.2	1223	28 US-09-705-256A-6006	Sequence 6006, Ap
15	1065.6	52.2	1223	60 US-60-164-285-6006	Sequence 6006, Ap
16	1051.2	51.5	1203	80 US-60-360-207-2119	Sequence 2119, Ap
17	1023.8	50.2	1041	16 US-09-205-070-8140	Sequence 8140, Ap
18	1023.8	50.2	1041	17 US-09-340-623-8140	Sequence 8140, Ap
19	1023.8	50.2	1041	33 US-09-898-888-8140	Sequence 8140, Ap
20	1023.8	50.2	1041	33 US-09-898-888A-8140	Sequence 8140, Ap
21	1012.8	49.6	1039	17 US-09-359-922-3465	Sequence 3465, Ap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

XX	Claim 25; SEQ ID No 1507; 487pp; English.
PS	
VV	

CC The present invention relates to human single exon nucleic acid probes  
CC (SENPs). The present sequence is one such probe. The SENPs are derived  
CC from human Hela cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [ftp://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

50 Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;

Query Match	Score	DB	Length
Best local similarity	7.18	145	22
Best local similarity	100.08	22	145

Best Local Similarity 100.08; Pred. No. 1.4e-29;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

183 AGGCTATGAGGCCCTAGACATCATCACCAAGTTTCTGCCCAAGAGCAGAGATCT 242

Db 450 AGGCTATGAAGCCTAGACATCATCACCAACAAGTTTCTGCCCAAGAGCAGAGAATCT 391

243 GCCGGCACCATGATCAGCTTTGTGGATCCTCTTGTGACCAATTACACAGTGTGGACT 302

Db 390 GCCGGCACCACATGATCAGCTTGTGGATCCTCTGTGACCAATTACACAGTGGTGGACT 331

QY 303 TCAGAAATAGAGCAACTGCTCTGAT 327

Db 330 TCAGAAATAGAGCAACTGCTCTGAT 306

```
Search completed: April 21, 2003, 13:37:09
Job time : 568.42 secs
```

```
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 1510; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;
XX
Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.4e-29;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 183 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 242
DB 450 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 391
QY 243 GCCGGCACCACATGATCAGCTTTGTGGATCCTTGTGACCAATTACACAGTGGTGGACT 302
DB 390 GCCGGCACCACATGATCAGCTTTGTGGATCCTTGTGACCAATTACACAGTGGTGGACT 331
QY 303 TCAGAAATAGAGCAACTGCTGTGAT 327
DB 330 TCAGAAATAGAGCAACTGCTGTGAT 306
XX
RESULT 14
AAK26966/C
ID AAK26966 standard; DNA; 457 BP.
XX
AC AAK26966;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 1523.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PT Penn SG, Hanzel DK, Chen W, Rank DR;
```

```
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 1523; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other;
XX
Query Match 7.1%; Score 145; DB 22; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.4e-29;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 183 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 242
DB 450 AGGCTATGAAGGCGCTAGACATCATCACCACACAGGTTTCTGCCAGAGCAGAGAATCT 391
QY 243 GCCGGCACCACATGATCAGCTTTGTGGATCCTTGTGACCAATTACACAGTGGTGGACT 302
DB 390 GCCGGCACCACATGATCAGCTTTGTGGATCCTTGTGACCAATTACACAGTGGTGGACT 331
QY 303 TCAGAAATAGAGCAACTGCTGTGAT 327
DB 330 TCAGAAATAGAGCAACTGCTGTGAT 306
XX
RESULT 15
AA11574/C
ID AA11574 standard; DNA; 457 BP.
XX
AC AA11574;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #1507 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
```

PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-496933/54.  
 XX  
 PT New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes  
 PS  
 PS Claim 1; SEQ ID NO 1534; 327bp + sequence listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and Bt 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for  
 CC verifying the expression of regions of genomic DNA predicted to  
 CC encode proteins. They are useful for gene discovery, and for  
 CC determining predisposition and/or prognosing breast disease. Gene  
 CC expression analysis is useful for assessing the toxicity of chemical  
 CC agents on cells. The microarray of this invention presents a far greater  
 CC diversity of probes for measuring gene expression, with far less bias  
 CC than expressed sequence tag microarrays. The method is suitable for  
 CC rapid production of functional information from genomic sequence. The  
 CC present sequence is a single exon nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other:  
 Query Match 7.1%; Score 145; DB 22; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-29;  
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 183 AGGCTATGAGAGGCTAGACATCATCAACAAGTTTCTGCCAAGAGCAGAAATCT 242  
 DB 450 AGGCTATGAGAGGCTAGACATCATCAACAAGTTTCTGCCAAGAGCAGAAATCT 391  
 OY 243 GCCGGCACACATGATGAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACT 302  
 DB 390 GCCGGCACACATGATGAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACT 331  
 OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327  
 DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306  
 RESULT 12  
 ABA53262/c  
 ID ABA53262 standard; DNA: 457 BP.  
 XX  
 AC ABA53262;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human foetal liver single exon nucleic acid probe #1567.  
 XX  
 KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00669.  
 XX

PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632386.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-483447/52.  
 XX  
 PT human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 PS  
 PS Claim 1; SEQ ID NO 1567; 639bp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX  
 SQ Sequence 457 BP; 120 A; 102 C; 110 G; 125 T; 0 other:  
 Query Match 7.1%; Score 145; DB 22; Length 457;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-29;  
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 183 AGGCTATGAGAGGCTAGACATCATCAACAAGTTTCTGCCAAGAGCAGAAATCT 242  
 DB 450 AGGCTATGAGAGGCTAGACATCATCAACAAGTTTCTGCCAAGAGCAGAAATCT 391  
 OY 243 GCCGGCACACATGATGAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACT 302  
 DB 390 GCCGGCACACATGATGAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACT 331  
 OY 303 TCAGAAATAGAGCAACTGCTCTGAT 327  
 DB 330 TCAGAAATAGAGCAACTGCTCTGAT 306  
 RESULT 13  
 AAK01519/c  
 ID AAK01519 standard; DNA: 457 BP.  
 XX  
 AC AAK01519;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 1510.  
 XX  
 KM Human; brain expressed exon; gene expression analysis; probe;  
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KM epilepsy; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 XX



KM pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
OS WO200160860-A2.  
PN 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US05171.  
PF 17-FEB-2000; 2000US-183319P.  
XX 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
XX 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
XX 13-DEC-2000; 2000US-255281P.  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
PI WPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX Claim 1; Page 3210; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX Sequence 545 BP; 179 A; 94 C; 91 G; 144 T; 37 other;  
SQ  
Query Match 8.0%; Score 163; DB 23; Length 545;  
Best Local Similarity 79.1%; Pred. No. 1.7e-34;  
Matches 193; Conservative 0; Mismatches 49; Indels 2; Gaps 1;  
QY 1800 ATGCAGAGAAAGGCGCATCTCCATTTAGATGATTAAGCAACCAACTAGTTCGGAAT 1859  
DB 20 AAGCGGGNANAGCCATCTCCATTTAGATGATTAAGCAACCAACTAGTTCGGAAT 79  
QY 1860 TCTACAGAGAGAGGAGATCAGACTGAGAGAGCTGACATAGACTTGAAGCCAAAG 1919  
DB 80 TTTTACACAGAGAGGAGATTAACCTGAGAGAGCTGACATAGACTTGAAGCCAAAG 139  
QY 1920 ACTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTGAG 1979  
DB 140 ACTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTGAG 199  
QY 1980 TTAC-AAACTATATTTTATTTAGATTAAATTAAGAAAAATTTACAGAAAAAAA 2037  
DB 200 TTACAAATTTAGATTTTATTTAGAGTTGATTAAGAAAACTGAACGACCATCAAC 259  
QY 2038 AAAA 2041  
DB 260 AAAA 263  
RESULT 9

ABV49413  
ID ABV49413 standard; CDNA: 582 BP.  
XX AC ABV49413;  
XX 17-SEP-2002 (first entry)  
XX Human prostate expression marker CDNA 49404.  
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KM pharmacogenomic marker; gene; ss.  
XX Homo sapiens.  
XX WO200160860-A2.  
XX 23-AUG-2001.  
XX 20-FEB-2001; 2001WO-US05171.  
XX 17-FEB-2000; 2000US-183319P.  
XX 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
XX 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
XX 13-DEC-2000; 2000US-255281P.  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX Schlegel R, Endege WO, Monahan JE;  
PI WPI; 2001-662795/76.  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX Claim 1; Page 9652; 11750pp; English.  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX Sequence 582 BP; 201 A; 114 C; 111 G; 156 T; 0 other;  
SQ  
Query Match 7.9%; Score 162; DB 23; Length 582;  
Best Local Similarity 84.7%; Pred. No. 3.3e-34;  
Matches 205; Conservative 0; Mismatches 35; Indels 2; Gaps 2;  
QY 1800 ATGCAGAGAAAGGCGCATCTCCATTTAGATGATTAAGCAACCAACTAGTTCGGAAT 1859  
DB 61 AAGCGGGNANAGCCATCTCCATTTAGATGATTAAGCAACCAACTAGTTCGGAAT 120  
QY 1860 TCTACAGAGAGAGGAGATCAGACTGAGAGAGCTGACATAGACTTGAAGCCAAAG 1919  
DB 121 TCTACACAGAGAGGAGATGAGACTGAGAGAGCTGACATAGACTTGAAGCCAAAG 180  
QY 1920 ACTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTGAG 1978  
DB 181 ACTTGAATTTGGAGCTGCTCATGTGTGATTAATATCACTGCTCTTTCTATTGAG 240



expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in AB161664 to AB170110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. MI can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of MI, and the data is sufficient to convey the chemical structure and/or properties of the agent. MI can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms' tumour.

Sequence 355 BP; 112 A; 74 C; 54 G; 115 T; 0 other:

Query Match 16.8%; Score 342.6; DB 24; Length 355;  
Best Local Similarity 98.9%; Pred. No. 4e-84;  
Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1684 CCTTAAAGAGTTTATGTCCTGACCTGCTAAATTTCTATTTCCAGATGCTT 1743  
DB 349 CCGTTAAAGAGTTTATGTCCTGACCTGCTAAATTTCTATTTCCAGATGCTT 290  
QY 1744 TTGTAGATGACGAGATATTTGTGAGCCCATATTGGAGTCTGATTGAGTGAATGG 1803  
DB 289 TTGTAGATGACGAGATATTTGTGAGCCCATATTGGAGTCTGATTGAGTGAATGG 230  
QY 1804 CAGGAAGGGCCATCTCCATTCAGATGATTAAGTGAACCAACATGTTCCGAATCTTA 1863  
DB 229 CAGGAAGGGCCATCTCCATTCAGATGATTAAGTGAACCAACATGTTCCGAATCTTA 170  
QY 1864 CAGGAAGGGAGGATTCAGACTGAGGAAGCTGTGACATGAGCTTGAACCAAGACTT 1923  
DB 169 CAGGAAGGGAGGATTCAGACTGAGGAAGCTGTGACATGAGCTTGAACCAAGACTT 110  
QY 1924 TGAATTTGCGAGCTGCTCATGCTGAGTTTATTCACAGCTGCTTTTATGAGTTAC 1983  
DB 109 TGAATTTGCGAGCTGCTCATGCTGAGTTTATTCACAGCTGCTTTTATGAGTTAC 50  
QY 1984 AAATCTATATTTTATGAAAGTTTAAATTAAGAAAAAATTTTACAGAAA 2032  
DB 49 AAATCTATATTTTATGAAAGTTTAAATTAAGAAAAAATTTTACAGAAA 1

RESULT 7  
AAA0911  
ID AAA0911 standard; cDNA; 300 BP.

AC AAA0911;  
XX  
DT 19-MAY-2000 (first entry)  
XX  
DE Human colon cancer cell line polynucleotide sequence SEQ ID NO:902.  
XX  
KW Human: colon cancer; tumour; diagnosis; gene expression product;  
KW probe; detection; cancerous state; metastasis; identification;  
KW breast cancer; oestrogen receptor-positive breast cancer; therapy;  
KW oestrogen receptor-negative breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
PN MO9958675-A2.  
XX  
PD 18-NOV-1999.  
XX  
XX  
PF 13-MAY-1999; 99WO-US10602.  
XX  
PR 14-MAY-1998; 98US-0085426.  
PR 15-MAY-1998; 98US-0085537.  
PR 15-MAY-1998; 98US-0085696.  
PR 21-OCT-1998; 98US-0105234.

PR 27-OCT-1998; 98US-0105877.  
XX  
XX (CHIR) CHIRON CORP.  
PA (HSE-) HSE INC.  
XX  
XX Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J,  
PI Reinhard C, Giese K, Randazzo F, Kennedy GC, Pot D, Kassam A,  
PI Lamson G, Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I,  
PI Leshkowitz D, Kita D, Garcia V, Jones LW, Stache-Crain B;  
XX  
XX WPI; 2000-126369/11.

Polynucleotide library used to determine cancerous states of mammalian cells -  
PS Claim 1; Page 399; 1097pp; English.

CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
CC libraries constructed from human colon cancer cell lines. The present  
CC invention also describes a method of detecting differentially expressed  
CC genes correlated with a cancerous state of a mammalian cell, comprising  
CC detecting at least one differentially expressed gene product in a test  
CC sample derived from a cell suspected of being cancerous, where detection  
CC of the differentially expressed gene product is correlated with a  
CC cancerous state of the cell from which the test sample was derived.  
CC The polynucleotide sequences can be used in a method for detecting  
CC differentially expressed genes correlated with a cancerous state of a  
CC mammalian cell. The polynucleotides can also be used as probes for  
CC detecting and mapping related genes. They can be used in diagnosis and  
CC prognosis of diseases and disorders (e.g. identification of  
CC pre-metastatic or metastatic cancerous states, stages of cancer, or  
CC responsiveness of cancer to therapy). This is particularly for breast  
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
CC negative breast cancer, lung cancer, and colon cancer.  
XX  
SQ Sequence 300 BP; 83 A; 69 C; 72 G; 76 T; 0 other;

Query Match 11.1%; Score 226.8; DB 21; Length 300;  
Best Local Similarity 95.1%; Pred. No. 3.1e-52;  
Matches 234; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 919 TAACAGAGCTTCTAAAGAAAGACAGCTGTCCTATGCTAGCTAGAGT 978  
DB 55 TAAAGACGTTTGTGAGAGAGACCTGGTCCATTTGCCCCCTGTATGAGAGT 114  
QY 979 ATCTGATGCTCGAAGTGGAGAGCTGTCTTGAACCTGCTTGAATCGTCAAG 1038  
DB 115 ATCTGATGCTCGAAGTGGAGAGCTGTCTTGAACCTGCTTGAATCGTCAAG 174  
QY 1039 TTTTCATCCAGGGCCACAGCCCTACAGCCCATTAAGATGCAATCAAGAGCTGA 1098  
DB 175 TTTTCATCCAGGGCCACAGCCCTACAGCCCATTAAGATGCAATCAAGAGCTGA 234  
QY 1099 GAACAAGAGAACTTATACCTGTGTGACCTGTGATGCAATCATTCATGGAGATCGGA 1158  
DB 235 GAACAAGAGAACTTATACCTGTGTGACCTGTGATGCAATCATTCATGGAGATCGGA 294  
QY 1159 ATGGGC 1164  
DB 295 ATGGGC 300

RESULT 8  
ABV19646  
ID ABV19646 standard; cDNA; 545 BP.  
XX  
AC ABV19646;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 19637.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

CC blood-related disorders (e.g. haemophilia), reproductive disorders  
 CC (e.g. infertility) and infectious disorders (e.g. Influenza). The  
 CC polynucleotides of the invention can also be used in gene therapy.  
 CC AA40785-AA51684 represent cDNA sequences encoding for the novel human  
 CC enzyme polypeptides of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.

XX Sequence 717 BP, 184 A; 177 C; 187 G; 162 T; 7 other;

## Query Match

Best Local Similarity 32.4%; Score 661.8; DB 22; Length 717;

Pred. No. 5,6e-172;

Matches 685; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

QY 1 CTGCCATAAGATGCGCTCCGTGGCGGTGACAGACAGTCTCTGTGGGAGTGGCTCAG 60  
 DB 1 CTGCCATAAGATGCGCTCCGTGGCGGTGACAGACAGTCTCTGTGGGAGTGGCTCAG 60  
 QY 61 GGGCGTGAAGAGGACCTCTCTGTAGTGTCTCGGGGCGACGGGCGCAATC 120  
 DB 61 GGGCGTGAAGAGGACCTCTCTGTAGTGTCTCGGGGCGACGGGCGCAATC 120  
 QY 121 CACGCTGGCGTTCAGCTAGCGCCAGCGGCTCGGAGATGCTCAGCGCTGACTCAT 180  
 DB 121 CACGCTGGCGTTCAGCTAGCGCCAGCGGCTCGGAGATGCTCAGCGCTGACTCAT 180  
 QY 181 GCAGGTCTATGAAGGCTTACATCATCACAACAAGTTTCTGCCCAAGCAGAGAA 240  
 DB 181 GCAGGTCTATGAAGGCTTACATCATCACAACAAGTTTCTGCCCAAGCAGAGAA 240  
 QY 241 CTGGCGGACCATGATAGTGTGTGATCTCTGTGACCAATTAACAGTGTGTA 300  
 DB 241 CTGGCGGACCATGATAGTGTGTGATCTCTGTGACCAATTAACAGTGTGTA 300  
 QY 301 CTTCAGAAATAGAGCAACTCTCTGATGAGATATTTTGGCCGAGCAAAATTCCTAT 360  
 DB 301 CTTCAGAAATAGAGCAACTCTCTGATGAGATATTTTGGCCGAGCAAAATTCCTAT 360  
 QY 361 TTTTGTGGAGAGCAATTTATTACATTGATCTGCTCTGGAAGTTCTGTCTATAC 420  
 DB 361 TTTTGTGGAGAGCAATTTATTACATTGATCTGCTCTGGAAGTTCTGTCTATAC 420  
 QY 421 CAAGCGCCGAGATGGGAGCTAGAGAAAGTGAATGACCAAAAGTGGAGCTTGAAGA 480  
 DB 421 CAAGCGCCGAGATGGGAGCTAGAGAAAGTGAATGACCAAAAGTGGAGCTTGAAGA 480  
 QY 481 GGATGGTCTTGTACTTCAACAAAGCCTAAGCAGGTGACCCAGAAATGGTCCAAAGT 540  
 DB 481 GGATGGTCTTGTACTTCAACAAAGCCTAAGCAGGTGACCCAGAAATGGTCCAAAGT 540  
 QY 541 GCATCCACTGACAAAGCAAGAGTGGCGAGCTTGAAGTTTGAAGAAACAGAGAT 600  
 DB 541 GCATCCACTGACAAAGCAAGAGTGGCGAGCTTGAAGTTTGAAGAAACAGAGAT 600  
 QY 601 CTCTCATAGTGAATTTCTCATGTCACATATA-CGGAAGAAGTGGTGTCCCC-TTGA 658  
 DB 601 CTCTCATAGTGAATTTCTCATGTCACATATA-CGGAAGAAGTGGTGTCCCC-TTGA 658  
 QY 659 GGTCTCTGAGAGTCTCTAACCCTTGACATCTTTGG 694  
 DB 661 GGTCTCTGAGAGTCTCTAACCCTTGACATCTTTGG 696

## RESULT 6

ABL68710/c  
 ID ABL68710 standard; DNA; 355 BP.

XX ABL68710;

DT 15-MAY-2002 (first entry)

XX Kidney cancer related gene sequence SEQ ID NO:7047.

XX Human: cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumor; carcinos;  
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KW gene; ds.  
 OS Homo sapiens.  
 XX WO200194629-A2.  
 PD 13-DEC-2001.  
 XX 30-MAY-2001; 2001WO-US10838.  
 PF 05-JUN-2000; 2000US-209473P.  
 PR 05-JUN-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 25-SEP-2000; 2000US-235280P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 27-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237295P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 (AVAL- ) AVALON PHARM.  
 XX Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;  
 PI Soppet DR, Weaver Z;  
 XX MPI: 2002-188264/24.  
 XX Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set  
 XX  
 PS Claim 1: SEQ ID 7047; 44pp; English.  
 XX The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in

PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226686.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 12-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0246617.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-465566/50.  
 P-PSDB; AAU23398.  
 DR Novel polypeptides and polynucleotides useful for diagnosing,  
 PT preventing, treating, neutral, immune system, muscular, reproductive,  
 PT pulmonary, cardiovascular, renal, proliferative disorders and cancerous  
 PT diseases  
 XX  
 PS Claim 4; SEQ ID No 494; 1180pp; English.  
 CC The present invention relates to the isolation of novel human enzyme  
 CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences  
 CC encoding them. The enzyme polypeptides of the invention may comprise the  
 CC functional classes of oxidoreductases, transferases, hydrolases, lyases,  
 CC isomerases or ligases. The sequences of the invention are useful in the  
 CC diagnosis, treatment, prevention and/or prognosis of a wide range of  
 CC disorders including hyperproliferative disorders (e.g. cancer),  
 CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders  
 CC (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),  
 CC metabolic disorders (e.g. phenylketonuria), inflammatory disorders  
 CC (e.g. asthma), cardiovascular disorders (e.g. atherosclerosis),

PR 17-NOV-2000; 2000US-249209P.  
 PR 17-NOV-2000; 2000US-249210P.  
 PR 17-NOV-2000; 2000US-249211P.  
 PR 17-NOV-2000; 2000US-249212P.  
 PR 17-NOV-2000; 2000US-249213P.  
 PR 17-NOV-2000; 2000US-249214P.  
 PR 17-NOV-2000; 2000US-249215P.  
 PR 17-NOV-2000; 2000US-249216P.  
 PR 17-NOV-2000; 2000US-249217P.  
 PR 17-NOV-2000; 2000US-249218P.  
 PR 17-NOV-2000; 2000US-249219P.  
 PR 17-NOV-2000; 2000US-249244P.  
 PR 17-NOV-2000; 2000US-249245P.  
 PR 17-NOV-2000; 2000US-249264P.  
 PR 17-NOV-2000; 2000US-249265P.  
 PR 17-NOV-2000; 2000US-249297P.  
 PR 17-NOV-2000; 2000US-249299P.  
 PR 17-NOV-2000; 2000US-249300P.  
 PR 01-DEC-2000; 2000US-250160P.  
 PR 01-DEC-2000; 2000US-250391P.  
 PR 05-DEC-2000; 2000US-251030P.  
 PR 05-DEC-2000; 2000US-251988P.  
 PR 05-DEC-2000; 2000US-251989P.  
 PR 06-DEC-2000; 2000US-251479P.  
 PR 08-DEC-2000; 2000US-251856P.  
 PR 08-DEC-2000; 2000US-251858P.  
 PR 08-DEC-2000; 2000US-251859P.  
 PR 08-DEC-2000; 2000US-251989P.  
 PR 08-DEC-2000; 2000US-251990P.  
 PR 11-DEC-2000; 2000US-254097P.  
 PR 05-JAN-2001; 2001US-259678P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-476161/51.  
 XX P-PSDB; ABB10476.  
 PT Isolated nucleic acid molecule encoding an inflammation-associated  
 PT polypeptide is used in preventing, treating or ameliorating a medical  
 PT condition -  
 XX  
 XX

Claim 1; SEQ ID NO: 364; 859pp + Sequence Listing; English.

CC The present invention provides human cDNAs, proteins and related genomic  
 CC DNAs. These can be used in the treatment of neural, immune system,  
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,  
 CC renal and proliferative disorders and inflammation. The present sequence  
 CC is a cDNA of the invention.

SO Sequence 717 BP; 184 A; 177 C; 187 G; 162 T; 7 other;

Query Match 32.4%; Score 661.8; DB 22; Length 717;  
 Best Local Similarity 98.4%; Pred. No. 5,6e-172;  
 Matches 685; Conservative 2; Mismatches 7; Indels 2; Gaps 2;

QY 1 CTGCGATGAAGTGGCGTCCGTCGCGGCTGCGAGCAGAGTTCCTGTGGCAGTGGCTCAG 60  
 DB 1 CTGCGATGAAGTGGCGTCCGTCGCGGCTGCGAGCAGAGTTCCTGTGGCAGTGGCTCAG 60  
 QY 61 GGGCGTCAAGGACCTTACTCTTGTAGTATCTCTGGGGCCACGGGCGCAGCAATC 120  
 DB 61 GGGCGTCAAGGACCTTACTCTTGTAGTATCTCTGGGGCCACGGGCGCAGCAATC 120  
 QY 121 CAGCGTGGCTGTGAGCTAGCGCAGCGCTGCGGCTGAGATCGTCAGCGCTGACTCAT 180  
 DB 121 CAGCGTGGCTGTGAGCTAGCGCAGCGCTGCGGCTGAGATCGTCAGCGCTGACTCAT 180  
 QY 181 GCAAGTGTATGAGGCTTACATCATCACCAACAGGTTCTGCCAAGAGCAGAGAT 240  
 DB 181 GCAAGTGTATGAGGCTTACATCATCACCAACAGGTTCTGCCAAGAGCAGAGAT 240  
 QY 241 CTGGCGGACACCATGATGAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGGTGA 300

DB 241 CTGGCGGACACCATGATGAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGGTGA 300  
 QY 301 CTTTCAGAAATAGAGCAACTGCTCTGATTTGAATATATTTGCCCGAGCAAAATTCCTAT 360  
 DB 301 CTTTCAGAAATAGAGCAACTGCTCTGATTTGAATATATTTGCCCGAGCAAAATTCCTAT 360  
 QY 361 TGTGTGGAGGAGCAACCAATTATTAATGTAATCTGCTCTGGAAGTTCTTGTCAATAC 420  
 DB 361 TGTGTGGAGGAGCAACCAATTATTAATGTAATCTGCTCTGGAAGTTCTTGTCAATAC 420  
 QY 421 CAAGCCGAGAGGATGGGCACTGAGAAAGTATTTGACCAAAAGTGGAGCTTGAAGA 480  
 DB 421 CAAGCCGAGAGGATGGGCACTGAGAAAGTATTTGACCAAAAGTGGAGCTTGAAGA 480  
 QY 481 GGATGCTCTTGTACTTCAACAAAGGCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGT 540  
 DB 481 GGATGCTCTTGTACTTCAACAAAGGCTTAAGCCAGGTGAGCCAGAAATGGCTGCCAAGT 540  
 QY 541 GCATCCACATGACAAACCAAGTGCCAGAGAGCTTGCAGATTGTAAGAAACAGAAAT 600  
 DB 541 GCATCCACATGACAAACCAAGTGCCAGAGAGCTTGCAGATTGTAAGAAACAGAAAT 600  
 QY 601 CTTCTATAGTGAATTTCTCCATCGTCACATTA-CGGAAGAGGTGGTGGTCCC-TTGA 658  
 DB 601 CTTCTATAGTGAATTTCTCCATCGTCACATTA-CGGAAGAGGTGGTGGTCCC-TTGA 658  
 QY 659 GGTCTCTGAAGTCTCTAACCCCTTGCAATCTTTGG 694  
 DB 659 GGTCTCTGAAGTCTCTAACCCCTTGCAATCTTTGG 694  
 DB 661 GGTCTCTGAAGTCTCTAACCCCTTGCAATCTTTGG 696

RESULT 5  
 ID AAS41268  
 AAS41268 standard; cDNA: 717 BP.

AC AAS41268;  
 DT 17-DEC-2001 (first entry)  
 XX  
 DE cDNA encoding novel human enzyme polypeptide #484.

XX Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;  
 KW ligase; hyperproliferative disorder; immunodeficiency disorder;  
 KW autoimmune disorder; neurological disorder; metabolic disorder;  
 KW inflammatory disorder; cardiovascular disorder; reproductive disorder;  
 KW blood-related disorder; infectious disorder; gene therapy; cytostatic;  
 KW anti arthritic; nephrotoxic; anticoagulant; ss.

OS Homo sapiens.

PN WO20015301-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01239.

XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216680.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217486.  
 PR 14-JUL-2000; 2000US-0218290.

QY 481 GGATGTCCTGCTACTGCACAAAGCCCTAAGCCAGGTGGACCCAGAAATGCTGCCAAGCT 540  
Db 481 GGATGTCCTGCTACTGCACAAAGCCCTAAGCCAGGTGGACCCAGAAATGCTGCCAAGCT 540  
QY 541 GCATCCCATGACGACAAAGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAGAGCAAT 600  
Db 541 GCATCCCATGACGACAAAGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAGAGCAAT 600  
QY 601 CTCCTCAGTGAATTTCTCCATGCTCAACATA-CGGAAGAAGGTGGTGTCCC-TTGA 658  
Db 601 CTCCTCAGTGAATTTCTCCATGCTCAACATA-CGGAAGAAGGTGGTGTCCC-TTGA 660  
QY 659 GGTCCTCTGAAGTCTCTCAACCTTGATCCTTTGG 694  
Db 661 GGTCCTCTGAAGTCTCTCAACCTTGATCCTTTGG 696

RESULT 4  
ABA06698  
ID ABA06698 standard; cDNA; 717 BP.  
XX ABA06698;  
AC ABA06698;  
XX 10-JAN-2002 (first entry)  
DT 10-JAN-2002 (first entry)  
XX  
DE Human cDNA SEQ ID NO: 364.  
XX  
KW Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200154474-A2.  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01349.  
XX  
PR 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 24-FEB-2000; 2000US-184664P.  
PR 02-MAR-2000; 2000US-186350P.  
PR 16-MAR-2000; 2000US-189874P.  
PR 17-MAR-2000; 2000US-190076P.  
PR 18-APR-2000; 2000US-198123P.  
PR 19-MAY-2000; 2000US-205515P.  
PR 07-JUN-2000; 2000US-209467P.  
PR 28-JUN-2000; 2000US-214886P.  
PR 30-JUN-2000; 2000US-215135P.  
PR 07-JUL-2000; 2000US-21647P.  
PR 07-JUL-2000; 2000US-216880P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225213P.  
PR 14-AUG-2000; 2000US-225214P.  
PR 14-AUG-2000; 2000US-225266P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225447P.  
PR 14-AUG-2000; 2000US-225757P.  
PR 14-AUG-2000; 2000US-225758P.  
PR 14-AUG-2000; 2000US-225759P.  
PR 18-AUG-2000; 2000US-226279P.  
PR 22-AUG-2000; 2000US-226881P.  
PR 22-AUG-2000; 2000US-226886P.

PR 22-AUG-2000; 2000US-227183P.  
PR 23-AUG-2000; 2000US-227009P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-229287P.  
PR 01-SEP-2000; 2000US-229343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 06-SEP-2000; 2000US-230437P.  
PR 06-SEP-2000; 2000US-230438P.  
PR 08-SEP-2000; 2000US-231242P.  
PR 08-SEP-2000; 2000US-231243P.  
PR 08-SEP-2000; 2000US-231244P.  
PR 08-SEP-2000; 2000US-231413P.  
PR 08-SEP-2000; 2000US-231414P.  
PR 08-SEP-2000; 2000US-232080P.  
PR 08-SEP-2000; 2000US-232081P.  
PR 12-SEP-2000; 2000US-231968P.  
PR 14-SEP-2000; 2000US-232397P.  
PR 14-SEP-2000; 2000US-232398P.  
PR 14-SEP-2000; 2000US-232399P.  
PR 14-SEP-2000; 2000US-232400P.  
PR 14-SEP-2000; 2000US-232401P.  
PR 14-SEP-2000; 2000US-233063P.  
PR 14-SEP-2000; 2000US-233064P.  
PR 14-SEP-2000; 2000US-233065P.  
PR 21-SEP-2000; 2000US-234223P.  
PR 21-SEP-2000; 2000US-234274P.  
PR 25-SEP-2000; 2000US-234997P.  
PR 25-SEP-2000; 2000US-234998P.  
PR 26-SEP-2000; 2000US-235484P.  
PR 27-SEP-2000; 2000US-235834P.  
PR 27-SEP-2000; 2000US-235836P.  
PR 29-SEP-2000; 2000US-236327P.  
PR 29-SEP-2000; 2000US-236367P.  
PR 29-SEP-2000; 2000US-236368P.  
PR 29-SEP-2000; 2000US-236369P.  
PR 29-SEP-2000; 2000US-236370P.  
PR 02-OCT-2000; 2000US-236802P.  
PR 02-OCT-2000; 2000US-237037P.  
PR 02-OCT-2000; 2000US-237038P.  
PR 02-OCT-2000; 2000US-237039P.  
PR 02-OCT-2000; 2000US-237040P.  
PR 13-OCT-2000; 2000US-238935P.  
PR 13-OCT-2000; 2000US-238937P.  
PR 20-OCT-2000; 2000US-240960P.  
PR 20-OCT-2000; 2000US-241221P.  
PR 20-OCT-2000; 2000US-241785P.  
PR 20-OCT-2000; 2000US-241786P.  
PR 20-OCT-2000; 2000US-241787P.  
PR 20-OCT-2000; 2000US-241808P.  
PR 20-OCT-2000; 2000US-241809P.  
PR 20-OCT-2000; 2000US-241826P.  
PR 01-NOV-2000; 2000US-244617P.  
PR 08-NOV-2000; 2000US-246474P.  
PR 08-NOV-2000; 2000US-246475P.  
PR 08-NOV-2000; 2000US-246476P.  
PR 08-NOV-2000; 2000US-246477P.  
PR 08-NOV-2000; 2000US-246478P.  
PR 08-NOV-2000; 2000US-246523P.  
PR 08-NOV-2000; 2000US-246524P.  
PR 08-NOV-2000; 2000US-246525P.  
PR 08-NOV-2000; 2000US-246526P.  
PR 08-NOV-2000; 2000US-246528P.  
PR 08-NOV-2000; 2000US-246529P.  
PR 08-NOV-2000; 2000US-246532P.  
PR 08-NOV-2000; 2000US-246609P.  
PR 08-NOV-2000; 2000US-246610P.  
PR 08-NOV-2000; 2000US-246611P.  
PR 08-NOV-2000; 2000US-246613P.  
PR 17-NOV-2000; 2000US-249207P.  
PR 17-NOV-2000; 2000US-249208P.

	PR	08-SEP-2000;	2000US-23114P.
	PR	08-SEP-2000;	2000US-23206P.
	PR	08-SEP-2000;	2000US-232081P.
	PR	12-SEP-2000;	2000US-231968P.
	PR	14-SEP-2000;	2000US-232397P.
	PR	14-SEP-2000;	2000US-232398P.
	PR	14-SEP-2000;	2000US-232399P.
	PR	14-SEP-2000;	2000US-232400P.
	PR	14-SEP-2000;	2000US-232401P.
	PR	14-SEP-2000;	2000US-233063P.
	PR	14-SEP-2000;	2000US-233065P.
	PR	21-SEP-2000;	2000US-234223P.
	PR	21-SEP-2000;	2000US-234274P.
	PR	25-SEP-2000;	2000US-234997P.
	PR	25-SEP-2000;	2000US-234998P.
	PR	26-SEP-2000;	2000US-235484P.
	PR	27-SEP-2000;	2000US-235834P.
	PR	27-SEP-2000;	2000US-235836P.
	PR	29-SEP-2000;	2000US-236327P.
	PR	29-SEP-2000;	2000US-236367P.
	PR	29-SEP-2000;	2000US-236368P.
	PR	29-SEP-2000;	2000US-236369P.
	PR	29-SEP-2000;	2000US-236370P.
	PR	02-OCT-2000;	2000US-236802P.
	PR	02-OCT-2000;	2000US-237037P.
	PR	02-OCT-2000;	2000US-237038P.
	PR	02-OCT-2000;	2000US-237039P.
	PR	02-OCT-2000;	2000US-237040P.
	PR	13-OCT-2000;	2000US-239935P.
	PR	13-OCT-2000;	2000US-239937P.
	PR	20-OCT-2000;	2000US-240960P.
	PR	20-OCT-2000;	2000US-241221P.
	PR	20-OCT-2000;	2000US-241785P.
	PR	20-OCT-2000;	2000US-241786P.
	PR	20-OCT-2000;	2000US-241787P.
	PR	20-OCT-2000;	2000US-241808P.
	PR	20-OCT-2000;	2000US-241809P.
	PR	20-OCT-2000;	2000US-241826P.
	PR	01-NOV-2000;	2000US-244617P.
	PR	08-NOV-2000;	2000US-246474P.
	PR	08-NOV-2000;	2000US-246475P.
	PR	08-NOV-2000;	2000US-246476P.
	PR	08-NOV-2000;	2000US-246477P.
	PR	08-NOV-2000;	2000US-246523P.
	PR	08-NOV-2000;	2000US-246524P.
	PR	08-NOV-2000;	2000US-246525P.
	PR	08-NOV-2000;	2000US-246526P.
	PR	08-NOV-2000;	2000US-246527P.
	PR	08-NOV-2000;	2000US-246528P.
	PR	08-NOV-2000;	2000US-246532P.
	PR	08-NOV-2000;	2000US-246609P.
	PR	08-NOV-2000;	2000US-246610P.
	PR	08-NOV-2000;	2000US-246611P.
	PR	17-NOV-2000;	2000US-249207P.
	PR	17-NOV-2000;	2000US-249208P.
	PR	17-NOV-2000;	2000US-249209P.
	PR	17-NOV-2000;	2000US-249210P.
	PR	17-NOV-2000;	2000US-249211P.
	PR	17-NOV-2000;	2000US-249212P.
	PR	17-NOV-2000;	2000US-249213P.
	PR	17-NOV-2000;	2000US-249214P.
	PR	17-NOV-2000;	2000US-249215P.
	PR	17-NOV-2000;	2000US-249216P.
	PR	17-NOV-2000;	2000US-249217P.
	PR	17-NOV-2000;	2000US-249218P.
	PR	17-NOV-2000;	2000US-249244P.
	PR	17-NOV-2000;	2000US-249245P.
	PR	17-NOV-2000;	2000US-249246P.
	PR	17-NOV-2000;	2000US-249265P.
	PR	17-NOV-2000;	2000US-249297P.
	PR	17-NOV-2000;	2000US-249299P.
	PR	17-NOV-2000;	2000US-249300P.
	PR	01-DEC-2000;	2000US-250160P.
	PR	01-DEC-2000;	2000US-250391P.
	PR	05-DEC-2000;	2000US-251030P.
	PR	05-DEC-2000;	2000US-251988P.
	PR	05-DEC-2000;	2000US-256719P.
	PR	06-DEC-2000;	2000US-251479P.
	PR	08-DEC-2000;	2000US-251856P.
	PR	08-DEC-2000;	2000US-251868P.
	PR	08-DEC-2000;	2000US-251869P.
	PR	08-DEC-2000;	2000US-251989P.
	PR	08-DEC-2000;	2000US-251990P.
	PR	11-DEC-2000;	2000US-254057P.
	PR	05-JAN-2001;	2001US-259678P.
	PA	(HUMA-) HUMAN GENOME SCI INC.	
	XX	Rosen CA, Barash SC, Ruben SM.	
	XX	WPI; 2001-476161/51.	
	DR	P-P5DB; ABB10278.	
	XX		

Db 1040 ATTGTCCTCCCTGCTATGCTTAGAGTATCTGATGCTGAACTGGGAGAGTCTGTT 1099  
QY 1010 CTGAACTGCTCTTGAAGTGTGCAAGTTTCATCCAGGGCCACAGCCTPACAGCACT 1069  
Db 1100 CTTGAACCTGCTTGAAGTGTGCAAGTTTCATCCAGGGCCACAGCCTPACAGCACT 1159  
QY 1070 CCATTAAGATGCAATGAAAGCTGAGAAACAGAAAGTATACCTGTGTACCTC 1129  
Db 1160 CCAATTAAGATGCAATGAAAGCTGAGAAACAGAAAGTATACCTGTGTACCTC 1219  
QY 1130 TGTGATCAATCATCATTTGGGGATCGCAATGGGAGCGCATTAATCCAAATCCAC 1189  
Db 1220 TGTGATCAATCATCATTTGGGGATCGCAATGGGAGCGCATTAATCCAAATCCAC 1279  
QY 1190 TTGAACCACTGAGAGAGAGAGAGATTTGAGCTCAGATGCTGTCAACACCATAGAAAGT 1249  
Db 1280 TTGAACCACTGAGAGAGAGAGAGATTTGAGCTCAGATGCTGTCAACACCATAGAAAGT 1339  
QY 1250 CAGAGTGTTCCTCCAGACTATACAAAGAACTTAAGGAGGATCCCCAGGCGAGAT 1309  
Db 1340 CAGAGTGTTCCTCCAGACTATACAAAGAACTTAAGGAGGATCCCCAGGCGAGAT 1399  
QY 1310 GATCAAGAGCTGAAGTGCAGCGTTTAAGACATGCTCCAGTGGCTTTGAAAGGTGGTG 1369  
Db 1400 GATCAAGAGCTGAAGTGCAGCGTTTAAGACATGCTCCAGTGGCTTTGAAAGGTGGTG 1459  
QY 1370 GGGATCCAGTTCAGAGAGGAGGGGTATTTGTCTCCAGTCTGGGCAAGAGTGCAT 1429  
Db 1460 GGGATCCAGTTCAGAGAGGAGGGGTATTTGTCTCCAGTCTGGGCAAGAGTGCAT 1519  
QY 1430 GCGGATCTCTGCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1489  
Db 1520 GCGGATCTCTGCATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1579  
QY 1490 CTCAGTCTCTATATATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1549  
Db 1580 CTCAGTCTCTATATATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1639  
QY 1550 TGGAAATATATAGTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1609  
Db 1640 TGGAAATATATAGTTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1699  
QY 1610 TTAAG 1669  
Db 1700 TTAAG 1759  
QY 1670 GGATGAGTGCATCCCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1729  
Db 1760 GGATGAGTGCATCCCTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1818  
QY 1730 ATTTCAGAGTCTTTTGTAGATGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1789  
Db 1819 ATTTCAGAGTCTTTTGTAGATGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878  
QY 1790 ATTGAGAGATGAG 1849  
Db 1879 ATTGAGAGATGAG 1938  
QY 1850 TTCTCGAATTTCTACAG 1909  
Db 1939 TTCTCGAATTTCTACAG 1998  
QY 1910 AAGACCAAGACTTTGAAATTTGGAGCTGCTCAGTGTGATTTATCTACTGCTGCT 1969  
Db 1999 AAGACCAAGACTTTGAAATTTGGAGCTGCTCAGTGTGATTTATCTACTGCTGCT 2058  
QY 1970 TTCTATTGAGTTACAATCTATATTTTATGAGTTTAAATTAAGAGAGAGAGAGAGAG 2029  
Db 2059 TTCTATTGAGTTACAATCTATATTTTATGAGTTTAAATTAAGAGAGAGAGAGAGAG 2118  
QY 2030 AAAAAAAAAAAAAA 2041  
Db 2119 AAAAAAAAAAAAAA 2130

RESULT 3  
AB06500  
ID AB06500 standard; cdna: 717 BP.  
XX  
AC AB06500;  
XX  
DT 10-JAN-2002 (first entry)  
XX  
DE Human cdna seq ID NO: 166.  
XX  
KW Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200154474-A2.  
PD  
XX 02-AUG-2001.  
PE 17-JAN-2001; 2001MO-US01349.  
XX  
PR 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 24-FEB-2000; 2000US-184664P.  
PR 02-MAR-2000; 2000US-186350P.  
PR 16-MAR-2000; 2000US-189874P.  
PR 17-MAR-2000; 2000US-190076P.  
PR 18-APR-2000; 2000US-198123P.  
PR 19-MAY-2000; 2000US-205515P.  
PR 07-JUN-2000; 2000US-209467P.  
PR 28-JUN-2000; 2000US-214866P.  
PR 30-JUN-2000; 2000US-215135P.  
PR 07-JUL-2000; 2000US-216647P.  
PR 07-JUL-2000; 2000US-216880P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225213P.  
PR 14-AUG-2000; 2000US-225214P.  
PR 14-AUG-2000; 2000US-225266P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225447P.  
PR 14-AUG-2000; 2000US-225577P.  
PR 14-AUG-2000; 2000US-225758P.  
PR 14-AUG-2000; 2000US-225759P.  
PR 18-AUG-2000; 2000US-226279P.  
PR 22-AUG-2000; 2000US-226811P.  
PR 22-AUG-2000; 2000US-226868P.  
PR 22-AUG-2000; 2000US-227182P.  
PR 23-AUG-2000; 2000US-227009P.  
PR 30-AUG-2000; 2000US-228924P.  
PR 01-SEP-2000; 2000US-228287P.  
PR 01-SEP-2000; 2000US-228343P.  
PR 01-SEP-2000; 2000US-229344P.  
PR 01-SEP-2000; 2000US-229345P.  
PR 05-SEP-2000; 2000US-229509P.  
PR 05-SEP-2000; 2000US-229513P.  
PR 06-SEP-2000; 2000US-230437P.  
PR 06-SEP-2000; 2000US-230438P.  
PR 08-SEP-2000; 2000US-231243P.  
PR 08-SEP-2000; 2000US-231243P.  
PR 08-SEP-2000; 2000US-231244P.  
PR 08-SEP-2000; 2000US-231413P.

|||||  
Db 1801 TGGCAGAGAAAGGCCATCTCCATTGATGATTAAGTGAACCAACTAGTCTCGAATT 1860  
Qy 1861 CTACAGAGAAAGGAGGAAATCAGACTGAGAGAGTGTGCACATAGACTTGAAGACCAAGA 1920  
Db 1861 CTACAGAGAAAGGAGGAAATCAGACTGAGAGAGTGTGCACATAGACTTGAAGACCAAGA 1920  
Qy 1921 CTTTGAATTTGGCAGCTGCTCATGTGTGAGTATTTATCACTGCTCTTCTATTGAGT 1980  
Db 1921 CTTTGAATTTGGCAGCTGCTCATGTGTGAGTATTTATCACTGCTCTTCTATTGAGT 1980  
Qy 1981 TACAATCTATATTTTATGAGTTTAAATAAAGAAAAAATTACAGAAAAAATAA 2040  
Db 1981 TACAATCTATATTTTATTTATTTAGATTAAATTAAGAAAAAATTACAGAAAAAATAA 2040  
Qy 2041 A 2041  
Db 2041 A 2041  
RESULT 2  
ABN59708  
ID ABN59708 standard; cDNA; 2130 BP.  
XX  
AC ABN59708;  
XX  
DT 28-JUN-2002 (first entry)  
XX  
DE Novel human coding sequence SEQ ID NO: 119.  
XX  
KW Human; antihaemetic; vulnery; antiinflammatory; immunomodulator;  
KW antileukemia; cerebroprotective; cytosolic; rheumatic; gene therapy;  
KW neuroprotective; antiparkinsonian; protein therapy; EST;  
XX expressed sequence tag; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200222660-A2.  
XX  
PD 21-MAR-2002.  
XX  
PE 10-SEP-2001; 2001WO-US26015.  
XX  
PR 11-SEP-2000; 2000US-0659671.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Dirmanac RT;  
XX  
DR WPI: 2002-292408/33.  
DR P-PSDB: ABB97295.  
XX  
PT An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis -  
XX  
PS Claim 1; SEQ ID NO 119; 509pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a coding sequence of the  
CC invention.  
XX  
SQ Sequence 2130 BP; 620 A; 439 C; 522 G; 549 T; 0 other;  
Query Match 94.1%; Score 1921.2; DB 24; Length 2130;

Best Local Similarity 95.9%; Pred. No. 0;  
Matches 2025; Conservative 0; Mismatches 8; Indels 79; Gaps 2;  
Qy 8 AAGATGGCGCTCGTGGCGGCTGCAGAGACATCTCTGTGGCAGTGGCTCAGGGCGCTG 67  
Db 20 ACGATTCGCTCCGTGGCGGCTGCAGAGACATCTCTGTGGCAGTGGCTCAGGGCGCTG 79  
Qy 68 CAACGGACCTTACCTCTTGTAGTATTCGCGGCGCAGGCGCACCAGGCAATTCACGCTG 127  
Db 80 CAACGGACCTTACCTCTTGTAGTATTCGCGGCGCAGGCGCACCAGGCAATTCACGCTG 139  
Qy 128 GCGTTGCGAGTACGAGCGCGGCTCGGCGGCTGAGATTCGACGCGTCACTCATGAGGCT 187  
Db 140 GCGTTGCGAGTACGAGCGCGGCTCGGCGGCTGAGATTCGACGCGTCACTCATGAGGCT 199  
Qy 188 TATGAGGCTTACATCATCATCACAACAGGTTTCTGCGCCAAAGACAGAGATCTGCCG 247  
Db 200 TATGAGGCTTACATCATCATCACAACAGGTTTCTGCGCCAAAGACAGAGATCTGCCG 259  
Qy 248 CACCATGATCAGCTTGTGATCCTCTGTGACCAATTACACAGTGGTGAATCAGA 307  
Db 260 CACCATGATCAGCTTGTGATCCTCTGTGACCAATTACACAGTGGTGAATCAGA 319  
Qy 308 AATAGAGCACTGCTCTGATTTGAGATATTTTGGCCGAGACAAATTCATTTGTTG 367  
Db 320 AATAGAGCACTGCTCTGATTTGAGATATTTTGGCCGAGACAAATTCATTTGTTG 379  
Qy 368 GGAGAACCAATTTATCATTTGATGATCTGCTGGAAGTTCTGTCAATACCAAGGCC 427  
Db 380 GGAGAACCAATTTATTCATTTGATGATCTGCTGGAAGTTCTGTCAATACCAAGGCC 439  
Qy 428 CAGGAGATGGGCACTGAGAAAGTGAATTCACGAAAGTGAAGCTTGAAGAGAGATGGT 487  
Db 440 CAGGAGATGGGCACTGAGAAAGTGAATTCACGAAAGTGAAGCTTGAAGAGAGATGGT 499  
Qy 488 CTGTACTTCAAAAGCGCTTAAGCGAGTGGACCCAGAAATGGCTGCCAAGTGCATCA 547  
Db 500 CTGTACTTCAAAAGCGCTTAAGCGAGTGGACCCAGAAATGGCTGCCAAGTGCATCA 559  
Qy 548 CATGACAAAGCAAGTGGCGAGAGCTTGCAGTTTGAAGAAACAGAAATCTCTCAT 607  
Db 560 CATGACAAAGCAAGTGGCGAGAGCTTGCAGTTTGAAGAAACAGAAATCTCTCAT 619  
Qy 608 AGTGAATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 667  
Db 620 AGTGAATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 679  
Qy 668 AAGTCTCTAACCCTTGCATCTTGGCTTATGCTGACACAGGAGTCTAGATAGGCG 727  
Db 680 AAGTCTCTAACCCTTGCATCTTGGCTTATGCTGACACAGGAGTCTAGATAGGCG 739  
Qy 728 TTGATTAAGAGGAGTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 787  
Db 740 TTGATTAAGAGGAGTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799  
Qy 788 CACAGACGCTATATCAGAGAAATGTTTGGAAATACCCAGAGACTATCAACATGATC 847  
Db 800 CACAGACGCTATATCAGAGAAATGTTTGGAAATACCCAGAGACTATCAACATGATC 859  
Qy 848 TTCCATATATTTGGCTTAAAGAAATTCACGAGTACCTGATCAGTGAAGAAATGACCA 907  
Db 860 TTCCATATATTTGGCTTAAAGAAATTCACGAGTACCTGATCAGTGAAGAAATGACCA 919  
Qy 908 CTGAGACTAGTAAACAGCTCTAAGAAAG-----GACTGCTGCTGCTGCTGCTGCTGCTGCT 938  
Db 920 CTGAGACTAGTAAACAGCTCTAAGAAAG-----GACTGCTGCTGCTGCTGCTGCTGCTGCT 949  
Qy 939 -----GACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949  
Db 980 AGATATGCGCGGAAACAAACCATGGTGTAAACCGTTTGTGAGCAGACAGCTGCTGCTGCT 1039  
Qy 950 ATTTGCCCCCTGTCTATGCTTAAAGATATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009



PT useful for enhancing longevity of a host and inhibiting tumour  
 PT formation

XX Claim 8; Fig 8; 93pp: English.

XX The present sequence represents the human homologue of gro-1, and  
 CC is referred to as hgro-1. The specification describes the five genes  
 CC of the Caenorhabditis elegans gro-1 operon (AA36071). The operon  
 CC contains the gro-1 gene (AA36072), the gro-1 gene (AA36074), the gro-2  
 CC gene (AA36075), the gro-2 gene (AA36075), and the hap-1 gene  
 CC (AA36077). The gro-1 gene can be used in a method for the diagnosis  
 CC and/or prognosis of cancer in a patient. Transgenic mice containing a  
 CC gene knock-out of a murine gene homologue of the gro-1 gene are useful  
 CC as models of aging and cancer. The proteins encoded by the genes are  
 CC useful for identifying compounds that affect the enzymatic activity  
 CC of these proteins. In order to enhance longevity of a host and inhibit  
 CC tumour formation, the gro-1 gene, together with the gro-1, gro-2,  
 CC gro-3 and hap-1 genes enables study of a physiological clock.

SQ Sequence 2041 BP; 589 A; 421 C; 502 G; 529 T; 0 other;

Query Match 100.0%; Score 2041; DB 20; Length 2041;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCATTAAGATGGCGTCCGTCGCGGCTGCAGCAGCAGTCTGTGGGAGTGGGCTCAG 60  
 DB 1 CTGCCATTAAGATGGCGTCCGTCGCGGCTGCAGCAGCAGTCTGTGGGAGTGGGCTCAG 60  
 QY 61 GGGCCATGAAGAGGAGCCCTTCTGTAGTATCTCGGGGCAACGGGCAACATC 120  
 DB 61 GGGCCATGAAGAGGAGCCCTTCTGTAGTATCTCGGGGCAACGGGCAACATC 120  
 QY 121 CAGCGTGGCGTTCACGTAGGCGGCGGCTCGGCGGAGATGCTGAGCCCTGACTCAT 180  
 DB 121 CAGCGTGGCGTTCACGTAGGCGGCGGCTCGGCGGAGATGCTGAGCCCTGACTCAT 180  
 QY 181 GCAGGTCTATGAAGGCGCTAGACATCATCACCACAAAGGTTCTGCCAAGAGCAGAAAT 240  
 DB 181 GCAGGTCTATGAAGGCGCTAGACATCATCACCACAAAGGTTCTGCCAAGAGCAGAAAT 240  
 QY 241 CTGCCGCGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 DB 241 CTGCCGCGACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 QY 301 CTTCAGAAATAGCAACATGCTCTGATGAAATATATTTGCCGAGACAAATTTCTAT 360  
 DB 301 CTTCAGAAATAGCAACATGCTCTGATGAAATATATTTGCCGAGACAAATTTCTAT 360  
 QY 361 TGTGTGGGAGGAAACCAAT 420  
 DB 361 TGTGTGGGAGGAAACCAAT 420  
 QY 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATGACGAAAGTGGAGCTTGAAGAAGA 480  
 DB 421 CAAGCCCCAGAGATGGGCACTGAGAAAGTATGACGAAAGTGGAGCTTGAAGAAGA 480  
 QY 481 GGAATGCTTGTACTTACAAAGCGCTTACAGGCTGAGCCAGAAATGGCTGCAAGCT 540  
 DB 481 GGAATGCTTGTACTTACAAAGCGCTTACAGGCTGAGCCAGAAATGGCTGCAAGCT 540  
 QY 541 GCATCCATGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 DB 541 GCATCCATGACAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 QY 601 CTCTCATAGTAATTTCTCATGCTTCAACATAGGAGAAAGTGGTGGTCCCTTGGAGG 660  
 DB 601 CTCTCATAGTAATTTCTCATGCTTCAACATAGGAGAAAGTGGTGGTCCCTTGGAGG 660  
 QY 661 TCCCTGAGAGTCTCTAACCCTTGACCTTTGGCTTCATGCTGACAGGAGGAGTTCAGA 720  
 DB 661 TCCCTGAGAGTCTCTAACCCTTGACCTTTGGCTTCATGCTGACAGGAGGAGTTCAGA 720

QY 721 TGAGCGCTTGATTAAGAGGAGTGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 DB 721 TGAGCGCTTGATTAAGAGGAGTGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
 QY 781 AGATTTTCACAGAGCGCTAT 840  
 DB 781 AGATTTTCACAGAGCGCTAT 840  
 QY 841 TGGTATCTTCCAAATCAATGCTTCAAGGAATTTTCAGAGTACCTGATCATGAGGAGAA 900  
 DB 841 TGGTATCTTCCAAATCAATGCTTCAAGGAATTTTCAGAGTACCTGATCATGAGGAGAA 900  
 QY 901 ATGCACACTGAGACACTAT 960  
 DB 901 ATGCACACTGAGACACTAT 960  
 QY 961 TGTCTATGCTTGAAGTATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1020  
 DB 961 TGTCTATGCTTGAAGTATCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1020  
 QY 1021 TCTTGAATTCGTCAAAAGTTTCATCCAGGCGCAAGGCTTACAGCAGCTCCAAATTAAGAT 1080  
 DB 1021 TCTTGAATTCGTCAAAAGTTTCATCCAGGCGCAAGGCTTACAGCAGCTCCAAATTAAGAT 1080  
 QY 1081 GCCATACATGAGCTGAGAACAAAGAGTATATACCTTGTGACCTCTGTATGAAAT 1140  
 DB 1081 GCCATACATGAGCTGAGAACAAAGAGTATATACCTTGTGACCTCTGTATGAAAT 1140  
 QY 1141 CATCATTTGGGAGTCCGAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 DB 1141 CATCATTTGGGAGTCCGAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200  
 QY 1201 GAAGAAAGAGAGAGATTTGAGTACAGTCTGATCAACACCATAGAGAGTGTGAGTGTTC 1260  
 DB 1201 GAAGAAAGAGAGAGATTTGAGTACAGTCTGATCAACACCATAGAGAGTGTGAGTGTTC 1260  
 QY 1261 CCCAGACTATTAACAAAGACCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 DB 1261 CCCAGACTATTAACAAAGACCTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 QY 1321 GAAATGAGGCTTTAAGAGACATGCTCAGTGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
 DB 1321 GAAATGAGGCTTTAAGAGACATGCTCAGTGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
 QY 1381 CAGAGGAGGAGGAGGATTTGCTCCAGCTGCGGCAAGAGAGTGTATGCGGATTC 1440  
 DB 1381 CAGAGGAGGAGGAGGATTTGCTCCAGCTGCGGCAAGAGAGTGTATGCGGATTC 1440  
 QY 1441 TGCATACAGAAAGCTCCACCATTTCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 DB 1441 TGCATACAGAAAGCTCCACCATTTCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 QY 1501 TATATATGAAGACAGAGGCTTGTGACCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 DB 1501 TATATATGAAGACAGAGGCTTGTGACCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 QY 1561 TAGTTCAAGGAAGCAATTTTCTTTTGAACCTTAAAGTCTATATATATATATATATATATATATATAT 1620  
 DB 1561 TAGTTCAAGGAAGCAATTTTCTTTTGAACCTTAAAGTCTATATATATATATATATATATATATATAT 1620  
 QY 1621 ACAGATTCACATTTTAT 1680  
 DB 1621 ACAGATTCACATTTTAT 1680  
 QY 1681 ATCCCTTTAAAGAGTTTATGCTGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1740  
 DB 1681 ATCCCTTTAAAGAGTTTATGCTGCTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1740  
 QY 1741 CTTTGTAGAT 1800  
 DB 1741 CTTTGTAGAT 1800  
 QY 1801 TGGCAGGAAGAGGAGCATCTCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:03:18 ; Search time 557.42 Seconds

(without alignments)  
8245.719 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041  
Sequence: 1 CTCGCATAGATGCGTCCG.....TTTACAGAAAAA 2041

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq.101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2041	100.0	2041	20	AAK36073	Human homologue of
2	1921.2	94.1	2130	24	ABN59708	Human homologue of
3	663	32.5	717	22	ABA06500	Human CDNA SEQ ID
4	661.8	32.4	717	22	ABA06698	Human CDNA SEQ ID
5	661.8	32.4	717	22	AA541268	CDNA encoding nove
6	342.6	16.8	355	24	AB168710	Kidney cancer rela
7	226.8	11.1	300	21	AAA00911	Human colon cancer
8	163	8.0	545	23	ABV19646	Human prostate exp
9	162	7.9	582	23	ABV49413	Human prostate exp

C	10	160.8	7.9	6399	23	ABL17248	Drosophila melanog
C	11	145	7.1	457	22	ABA42839	Human breast cell
C	12	145	7.1	457	22	ABA53262	Human foetal liver
C	13	145	7.1	457	22	AAK01519	Human brain expres
C	14	145	7.1	457	22	AAK26966	Human bone marrow
C	15	145	7.1	457	22	AAI11574	Probe #1507 for ge
C	16	145	7.1	457	22	AAI32862	Human breast cell
C	17	138	6.8	138	22	ABA47966	Human foetal liver
C	18	138	6.8	138	22	ABA65854	Human brain expres
C	19	138	6.8	138	22	AAK43265	Human bone marrow
C	20	138	6.8	138	22	AAK39928	Probe #10720 for g
C	21	138	6.8	138	22	AAI20787	Drosophila melanog
C	22	138	6.8	138	22	AAI46014	Drosophila melanog
C	23	75.8	3.7	6225	23	AB10424	C glutaminc codin
C	24	75.8	3.7	6225	23	AB10424	Arabidopsis thalia
C	25	75	3.7	903	22	AAH67098	Arabidopsis thalia
C	26	75	3.7	1026	22	AAH67881	Arabidopsis thalia
C	27	75	3.7	349980	22	AAH68530	Aspergillus oryzae
C	28	73	3.6	1238	21	AAH54043	S. epidermidis ope
C	29	71.4	3.5	1243	21	AAH54043	S. epidermidis ope
C	30	71.4	3.5	1713	21	AAH54043	Arabidopsis thalia
C	31	71.4	3.5	1713	21	AAH54043	Arabidopsis thalia
C	32	69.6	3.4	700	21	AAH12572	Arabidopsis thalia
C	33	65	3.2	582	24	ABN91871	Arabidopsis thalia
C	34	65	3.2	990	22	AAH53033	Arabidopsis thalia
C	35	65	3.2	3775	22	AAH54468	Arabidopsis thalia
C	36	62.6	3.1	1374	21	AAH49935	Arabidopsis thalia
C	37	61	3.0	1376	21	AAH37688	Arabidopsis thalia
C	38	60	2.9	60	24	ABN34729	Human spliced tran
C	39	58.4	2.9	4411529	22	AAI99682	Mycobacterium tube
C	40	57.6	2.8	918	24	ABO69238	Listeria monocytog
C	41	57.6	2.8	2944528	24	ABO69238	Listeria innocua D
C	42	52	2.5	1092	24	ABO69242	Listeria innocua C
C	43	52	2.5	495269	24	ABO67195	Enterococcus faeca
C	44	50.4	2.5	8519	20	AAI33043	Staphylococcus aur
C	45	49.8	2.4	989	18	AAV75111	

#### ALIGNMENTS

RESULT 1  
AAK36073 standard; DNA; 2041 BP.

AAK36073:

15-JUL-1999 (first entry)

Human homologue of gro-1, referred to as hgro-1.

gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene;

hap-1 gene; Cancer; aging; longevity; tumour formation;

physiological clock; ss.

Homo sapiens.

WO9910482-A1.

04-MAR-1999.

20-AUG-1998; 98WO-CA00803.

25-AUG-1997; 97CA-2210251.

(UYWC-) UNIV MCGILL.

Barnes T, Hekimi S, Lakowski B, Lemieux J;

WPI; 1999-190615/16.

Molecular identity of the gro-1 gene - useful for cancer diagnosis

and/or prognosis, and where compounds affecting encoded proteins are

Db 196247 GAGATCTTTTTTATGATGGTAATTCAGAAATTCAGTGCAT -CCTTTAAAGACCGTCTTT 196305

QY 1703 GTCCCTGACTCTGGCTAAATTTATCTATTTTCCAGATGCTTTTGTGATGACTGAAGTAT 1762

Db 196306 GTCCCTGGCCGTGGCTATAAATAATCCTGGTTTCCAGAGCTTTTGTGATGATGACTGAAGTAT 196365

QY 1763 TTGTAGACCACTATTTGGGAGTCTTAGATTTTGATGATGAATGGCAGGAAGGGCCATCTCCA 1822

Db 196366 TGTAAAGGCGAGTGAAGAGCTCTGCACTCGAATCGAAGACGGCNAAGAGATAG ------TGCA 196419

QY 1823 TTGAGATGATTTAACTGAACCAACTAGTTCTCGGAATTTTACAGAGAGAGGGAATCAG 1882

Db 196420 GAGGAGATGATTTAAACCAAACTCTCCACCTCTGTGAAT- ---CTACAGAAAGAGGGGCTCAG 196475

QY 1883 ACTGAGAGACGCTGACATATAGAGATGAAGAC -CAAGACCTTTGAAATTTGCGAGTGC 1941

Db 196476 GCTGAGGTGGTGTGGACCTGATCTTAAAGACATAAAGACACAGAGTCTCTAGCTGCT 19555

QY 1942 CATGTG 1947

Db 196536 CCTGTG 196541

Search completed: April 21, 2003, 15:52:47  
Job time : 7858.89 secs



QY	902	TGCACTGGAGACTAGTAACCACTTCTAAAGAAAGAACTGGTCCCATGTGCCCT	961
Db	228367	AAACAAAAACCGCTTTTGTAGACAGACAGCTGGGTTCACAGTCCCTCCACCGCCCCCCCC	228426
QY	962	GTCATGGCTTAAAGAGCTATCGATGTCGAAAGTGGAGAGTGTGCTTGAACCTGCT	1021
Db	228427	CCCCACCTCCCGCGCGTATTATGATTTTTCCAAGTGGAGAGTGTGCTTGTGAACCTGCT	228486
QY	1022	CTTGAATCGTGAAGATTTTCATCCAGGGCCACAAGCTACAGCCACTCAATAAAGATG	1081
Db	228487	CTCCATACAGTAACAAGTTTCATCCAGGGCTCAAACTTACAGCCACTCCGGTGAAGTG	228546
QY	1082	CCATATCAATGAAGCTGAGAACAAAGAAAGTATATCACTCTGTGACCTCTGTATGAATC	1141
Db	228547	CCGTACATGTAAC--AGAACAGAGAAAGTTGCCACATATGTGACCTTTGTGATTTGATC	228604
QY	1142	ATCATTTGGGGATTCGGGATTTGGGACGGCAGCAATAAATCCCAATCCCACTGAAACACTG	1201
Db	228605	ATCATTTGGGGACCAAGAAATGGGACGAC--	228632
QY	1202	AAGAAAGAAAGATTTGAGCTCAGATCTGTCAACACATPAGAAAGTCAGAGTGTTC	1261
Db	228633	-----ATTGACTTATGAGCTGTCAATGTGCTTGAAGAACCAAGTAAATCA	228678
QY	1282	CCAGACTATTAACAAGAACTTAAAGGAAAGGATCCCAAGGACAGATGATCAAGACTG	1321
Db	228679	CCAGACTGTGACTCGGGACGACAGAAAGGTG-----AGGGCACATGATCAAGAGCTC	228730
QY	1322	AAATGACACGTTTAAAGAACATATGCTCAGTGGCTTTGAAAGGTGGGATTCAGATTG	1381
Db	228731	AAGGCCAGTGTGTGGAACATG-----CCTTGCAGAGATGTGGGATTCAGAGCC	228781
QY	1382	AGAGAGGAGGGGTATGTTGTCTCCCA--GTCTGGGCAAGAGAGTCTATGCGAATTGTC	1440
Db	228782	AGAGAGGAGGGAGATGTGCTCTCCACAGCCTTGAGCTGAGAAAGCTGGGCAAGAGGCC	228841
QY	1441	TGCATAGCAGAAAGCTCCACACATTTCTTTGATGTGTTTAAAGTCTCAGTTCTC	1500
Db	228842	CGCCA-----TCCTTCATTTCTGTGTGTGG	228866
QY	1501	TATATTAAGAAACAGCAGGCTGTGACGCTCTGTGTGGCTGATGTGCTGAAATGATG	1560
Db	228867	TCTGCAATGGAAGACCGACAGCTCTTCAAGCCTTTGTGTGTGTGTGTGTGTATGATG	228926
QY	1561	TAGTTCAGAAAGCATTTTCTTTTCTTTGAACCTTAAAGGTTCTATTAATTAAGAACAC	1620
Db	228927	TAGTTC--AAGGGGATTTTCCCTTCAACCTTAAAGGTTTATTTTAAAGAGAGC	228985
QY	1621	ACAGATTCACATTTTATACATGAGGATCTTCTTTGTGTGTATATACAGGATTTGACTGC	1680
Db	228986	ACAGGTTCCACATCTTTC-----TTTATGCTGTCAACCAAGATTCAGATC	229033
QY	1681	ATCCCTTAAAGAAAGATTTTATGCTCCGACCTGCGCTGCTAAATTTATCTTCCAGATG	1740
Db	229034	ATCCCTTAAAGAGAGTCTTGTCCCTGCGCTGCTTAAATTAATCTATTTCCAGAGG	229093
QY	1741	CTTTGTATGATGACTGAAGTATTTGTGAGCCACATATTGGGAGTCTGATTTGAAGTAA	1800
Db	229094	CTTTGTATGATGACTGAAGTGTGTAAGGACACAGTCAAGAGTGTGAGACAGAGAAAGCG	229153
QY	1801	TGGCAGAAAGGCGCATCTCCATTTAGATGATTAAGTGAAACCAACTAGTCTCGCAATT	1860
Db	229154	AGGAAGAGAGTGTG-----CCGAGAGATGATTAAGCAGATTGCCACAGCTCTGTGAAT-	229206
QY	1861	CTACAGAAAGAGGAGATCAAGACGAGAAAGCTGTGACATAGGACTTGAACCAAGAA	1920
Db	229207	---CTACAGCAGAGGGGCTCAGGCTCAGGTGTGTGTGACCCAGATCTTGAAGACTTAAGA	229263

\*\*\* 35 unordered pieces.

ACCESSION AC093939

VERSION AC093939.4 GI:21723477

KEYWORDS HTG; HTGS; PHASE1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 239888)

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alshrocks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T., Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhan,C., Burck,P., Burckett,C., Burrell,K.L., Byrd,N.C., Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthett,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabril,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichteige,O., Lien,C., Liu,J., Liu,W., Lohseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogund,M., Okunolu,G., Otaguaye,N., Oyler,R., Pace,A., Payton,B., Peery,L., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umami,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE JOURNAL

REFERENCE 2 (bases 1 to 239888)

AUTHORS

Unpublished

TITLE JOURNAL

REFERENCE 3 (bases 1 to 239888)

AUTHORS

Unpublished

TITLE JOURNAL

REFERENCE 4 (bases 1 to 239888)

AUTHORS

Unpublished

TITLE JOURNAL

REFERENCE 5 (bases 1 to 239888)

AUTHORS

Unpublished

TITLE JOURNAL

REFERENCE 6 (bases 1 to 239888)

AUTHORS

Unpublished

TITLE JOURNAL

REFERENCE 7 (bases 1 to 239888)

AUTHORS

Unpublished

On Jul 10, 2002 this sequence version replaced g1:20335575.

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GEXO

Center clone name: CH230-172N14

#### Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 210507 bases at least Q40

Consensus quality: 213691 bases at least Q30

Consensus quality: 215781 bases at least Q20

NOTE: Estimated insert size may differ from sequence length (see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 35 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1	1220:	contig of 1220 bp in length
1221	1320:	gap of unknown length
1321	2370:	contig of 1050 bp in length
2371	2470:	gap of unknown length
2471	3511:	contig of 1041 bp in length
3512	3611:	gap of unknown length
3612	4645:	contig of 1034 bp in length
4646	4745:	gap of unknown length
4746	5993:	contig of 1248 bp in length
5994	6094:	gap of unknown length
6094	7118:	contig of 1025 bp in length
7119	7219:	gap of unknown length
7219	8598:	contig of 1381 bp in length
8600	8699:	gap of unknown length
8700	9998:	contig of 1299 bp in length
9999	10098:	gap of unknown length
10099	11207:	contig of 1109 bp in length
11208	11307:	gap of unknown length
11308	12832:	contig of 1525 bp in length
12833	12932:	gap of unknown length
12933	15681:	contig of 2749 bp in length
15682	15781:	gap of unknown length
15782	18002:	contig of 2221 bp in length
18003	18102:	gap of unknown length
18103	20292:	contig of 2190 bp in length
20293	20392:	gap of unknown length
20393	22189:	contig of 1797 bp in length
22190	22289:	gap of unknown length
22290	24818:	contig of 2529 bp in length
24819	24918:	gap of unknown length
24919	29103:	contig of 4185 bp in length
29104	29203:	gap of unknown length
29204	31883:	contig of 2680 bp in length
31884	31983:	gap of unknown length
31984	34715:	contig of 2732 bp in length
34716	34815:	gap of unknown length
34816	39740:	contig of 4925 bp in length
39741	39840:	gap of unknown length
39841	42402:	contig of 2562 bp in length
42403	42502:	gap of unknown length
42503	48347:	contig of 5845 bp in length
48348	48447:	gap of unknown length
48448	52407:	contig of 3960 bp in length
52408	52507:	gap of unknown length
52508	56832:	contig of 4325 bp in length
56833	56932:	gap of unknown length
56933	63189:	contig of 6257 bp in length
63190	63289:	gap of unknown length
63290	68850:	contig of 5561 bp in length
68851	68950:	gap of unknown length
68951	76141:	contig of 7191 bp in length
76142	76241:	gap of unknown length
76242	83247:	contig of 7006 bp in length
83248	83347:	gap of unknown length
83348	94069:	contig of 10722 bp in length

RESULT	13
AC093939	
LOCUS	
DEFINITION	Rattus norvegicus clone CH230-172N14, *** SEQUENCING IN PROGRESS
	AC093939 239888 bp DNA linear HTG 11-JUL-2002

```

Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Moser, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokwkw, S., Ogun, M., Okwona, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Swatek, A., Tabor, P., Tamerisa, A., Tamerisa, R., Tang, H.,
Tansay, D., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 181838)
Morley, K.C.
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 181838)
Morley, K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18139174.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GLTJ
Center Clone name: CH230-64124
----- Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 115969 bases at least Q40
Consensus quality: 122034 bases at least Q30
Consensus quality: 127422 bases at least Q20

-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1074 1073: contig of 1073 bp in length
1174 1173: gap of unknown length
2367 2367: contig of 1194 bp in length
2468 2467: gap of unknown length
3755 3755: contig of 1288 bp in length
3856 3855: gap of unknown length
5139 5138: contig of 1283 bp in length
6254 6254: gap of unknown length
7644 7644: gap of unknown length
9196 9196: contig of 1452 bp in length
9296 9296: gap of unknown length

9297 10423: contig of 1127 bp in length
10424 10523: gap of unknown length
10524 12093: contig of 1570 bp in length
12094 12193: gap of unknown length
12194 13606: contig of 1413 bp in length
13607 15139: gap of unknown length
15140 15239: gap of unknown length
15240 16519: gap of unknown length
16520 16630: gap of unknown length
16630 18029: contig of 1410 bp in length
18030 18129: gap of unknown length
18130 19537: gap of unknown length
19538 19637: gap of unknown length
19638 20773: contig of 1136 bp in length
20774 20873: gap of unknown length
20874 21940: contig of 1067 bp in length
21941 22040: gap of unknown length
22041 23977: contig of 1937 bp in length
23978 24077: gap of unknown length
24078 25464: contig of 1387 bp in length
25465 25564: gap of unknown length
25565 26881: contig of 1317 bp in length
26882 26981: gap of unknown length
26982 28875: gap of unknown length
28876 28975: gap of unknown length
28976 30545: contig of 1570 bp in length
30546 31672: contig of 1027 bp in length
31673 31772: gap of unknown length
31773 33195: contig of 1423 bp in length
33196 33295: gap of unknown length
33296 34682: contig of 1387 bp in length
34683 34782: gap of unknown length
34783 35966: contig of 1184 bp in length
35967 36066: gap of unknown length
36067 37755: contig of 1689 bp in length
37756 37855: gap of unknown length
37856 40161: contig of 2306 bp in length
40162 40261: gap of unknown length
40262 42177: contig of 1916 bp in length
42178 42277: gap of unknown length
42278 44344: contig of 2067 bp in length
44345 44444: gap of unknown length
44445 46472: contig of 2028 bp in length
46473 46572: gap of unknown length
46573 48054: contig of 1482 bp in length
48055 48154: gap of unknown length
48155 48390: contig of 1236 bp in length
48391 49490: gap of unknown length
49491 52986: contig of 3496 bp in length
52987 53086: gap of unknown length
53087 54647: contig of 1561 bp in length
54648 54747: gap of unknown length
54748 57112: contig of 2265 bp in length
57113 57212: gap of unknown length
57213 58488: contig of 1636 bp in length
58489 58948: gap of unknown length
58949 60817: contig of 1869 bp in length
60818 60917: gap of unknown length
60919 63443: contig of 2526 bp in length
63444 63543: gap of unknown length
63544 66294: contig of 2751 bp in length
66295 66394: gap of unknown length
66395 68140: contig of 1746 bp in length
68141 68241: gap of unknown length
68241 71198: contig of 2958 bp in length
71199 71298: gap of unknown length
71299 73990: contig of 2692 bp in length
73991 74090: gap of unknown length
74091 75872: contig of 1782 bp in length
75873 75972: gap of unknown length
75973 79069: contig of 3097 bp in length

```

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

TITLE

AUTHORS

TITLE

JOURNAL



```

misc_feature      vector_side:left
                  2379..22743
/note="assembly_fragment:01823
fragment_chain:1"
misc_feature      22844..45183
/note="assembly_fragment:01809
fragment_chain:2"
misc_feature      45284..49581
/note="assembly_fragment:01177
fragment_chain:2"
misc_feature      49682..58258
/note="assembly_fragment:00999
fragment_chain:2"
misc_feature      58359..64082
/note="assembly_fragment:01187
fragment_chain:2"
misc_feature      64183..72764
/note="assembly_fragment:00292
fragment_chain:2"
misc_feature      72865..86552
/note="assembly_fragment:00339
fragment_chain:3"
misc_feature      86553..93481
/note="assembly_fragment:01156
fragment_chain:3"
misc_feature      93582..101614
/note="assembly_fragment:01960
fragment_chain:3"
misc_feature      101715..107836
/note="assembly_fragment:00867
fragment_chain:4"
misc_feature      107937..144546
/note="assembly_fragment:01896
fragment_chain:4"
misc_feature      144647..155892
/note="assembly_fragment:00307"
misc_feature      155993..167875
/note="assembly_fragment:00528"
misc_feature      167976..183411
/note="assembly_fragment:01170"
misc_feature      183512..199863
/note="assembly_fragment:01894"
BASE COUNT      53282 a 45915 c 45981 g 53154 t 1531 others
ORIGIN
Query Match      41.1%; Score 839.4; DB 2: Length 199863;
Best Local Similarity 98.5%; Pred. No. 5.1e-208;
Matches 868; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 1151 GATCGCAATGGGCGACGACATTAATCCCAATCCCACTTGACCACTGAAGAAGA 1210
DB 26154 GATTTTGTCTTATCAGCGACATTAATCCCAATCCCACTTGACCACTGAAGAAGA 26095
QY 1211 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTAT 1270
DB 26094 AGAAGATTGGACTCAGATGCTGTCAACACCATAGAAAGTCAGAGTGTTCGCCAGACTAT 26035
QY 1271 AACCAAGAACCTTAAGGGAAGGATCCCGACGAGAAATGATCAGAGCTGAATTCACAC 1330
DB 26034 AACCAAGAACCTTAAGGGAAGGATCCCGACGAGAAATGATCAGAGCTGAATTCACAC 25975
QY 1331 GTTTAAGACATGTCACATGCTTTGGAAGAGTGTGGGATTCAGATTCAGAGAGAG 1390
DB 25974 GTTTAAGACATGTCACATGCTTTGGAAGAGTGTGGGATTCAGATTCAGAGAGAG 25915
QY 1391 GGGTATGTTTGTCTCCCACTGTGGCAAGAGAGTGTATGCGAATTCCTCTCATAGCAG 1450
DB 25914 GGGTATGTTTGTCTCCCACTGTGGCAAGAGAGTGTATGCGAATTCCTCTCATAGCAG 25855
QY 1451 AAAAGCTCCACCATTTTCTTTGATGTGTGTTTAAAGTCTACGCTCTCTATAAGAA 1510
DB 25854 AAAAGCTCCACCATTTTCTTTGATGTGTGTTTAAAGTCTACGCTCTCTATAAGAA 25795

```

```

QY 1511 ACAGCAGGCTGTGTCAGCTCCTGTGTGTGCTGATGTGTCTGGAATGATGATTCAGCA 1570
DB 25794 ACAGCAGGCTGTGTCAGCTCCTGTGTGTGCTGATGTGTCTGGAATGATGATTCAGCA 25735
QY 1571 AAGCATTTTTTTTTTCTTTGAAACCTTAAGGTTCTATATTAAAGCAGACAGATTCGA 1630
DB 25734 AAGCATTTTTTTTTTCTTTGAAACCTTAAGGTTCTATATTAAAGCAGACAGATTCGA 25675
QY 1631 CA-TTTTATACATGAGATCTCTGTTGGTGAATACAGAGATTGACTCCCTTTA 1689
DB 25674 CATTTTATACATGAGATCTCTGTTGGTGAATACAGAGATTGACTCCCTTTA 25616
QY 1690 AAAGAAGTTTATATGTCCTGACTGTGCTAAATATATCAATTTCCAGATGCTTTGAG 1749
DB 25615 AAAGAAGTTTATATGTCCTGACTGTGCTAAATATATCAATTTCCAGATGCTTTGAG 25556
QY 1750 ATGACTGAAGTATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTAATGACAGGA 1809
DB 25555 ATGACTGAAGTATTTGTGAGCCACATATTGGAGTTCTAGATTGAGTAATGACAGGA 25496
QY 1810 AGGGCCATCTCATTTGAGATGATTAAGTAAAGCAACATGATCTGGAATTTACAGAGA 1869
DB 25495 AGGGCCATCTCATTTGAGATGATTAAGTAAAGCAACATGATCTGGAATTTACAGAGA 25436
QY 1870 AGGAGGAATCAGACTGAGAGAGCTGTGACATAGACATTAAGACCAAGACTTTGAAT 1929
DB 25435 AGGAGGAATCAGACTGAGAGAGCTGTGACATAGACATTAAGACCAAGACTTTGAAT 25376
QY 1930 TTGCGAGCTGCTCATGCTGTGAGTTATATCACTGCTCTTTCTATGAGTTACAAATCT 1989
DB 25375 TTGCGAGCTGCTCATGCTGTGAGTTATATCACTGCTCTTTCTATGAGTTACAAATCT 25316
QY 1990 ATATTTTATGATGATTTAAATAAGAAAAATTTTACAGA 2030
DB 25315 ATATTTTATGATGATTTAAATAAGAAAAATTTTACAGA 25275
RESULT 12
AC106650 181838 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-64124, *** SEQUENCING IN PROGRESS
DEFINITION ***; 62 unordered pieces.
ACCESSION AC106650.2 GI:21734730
VERSION HTG: HTGS PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 181838)
Muzny,D.M., Adams,C., Adio-Oduola,B., All-oshman,F.R., Allen,C.,
Alsbrooks,S.L., Amaral-Lungue,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Blumage,K., Blumenthurg,K., Bonin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burnell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Datorre,S.R., David,R.,
Devilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Dublin,K.J.,
Einhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
Homs,J., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudh,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,J., Liu,W., Louisege,H.,

```

```

QY 1211 AGAAGATTGACACAGATGCTGTCACACACATAGAAAGTCAGATGTTCCCGACACTAT 1270
      |||
Db 93981 AGAAGATTGACACAGATGCTGTCACACACATAGAAAGTCAGATGTTCCCGACACTAT 1270
      |||
QY 1271 AACAAAGAACCTTAAGAGGAGGATCCCGACGAGATGATCAAGAGCTGAATGCAGC 1330
      |||
Db 93921 AACAAAGAACCTTAAGAGGAGGATCCCGACGAGATGATCAAGAGCTGAATGCAGC 1330
      |||
QY 1331 GTTTAAGACATGTCACAGTGCGCTTTGGAAGGTGGGATCCAGTTCCAGAGGAG 1390
      |||
Db 93861 GTTTAAGACATGTCACAGTGCGCTTTGGAAGGTGGGATCCAGTTCCAGAGGAG 1390
      |||
QY 1391 GGCTAATGTTGCTCCACAGTCTGGGCAAAAGATGCTATCCGGAATTCCTGCATACAG 1450
      |||
Db 93801 GGCTAATGTTGCTCCACAGTCTGGGCAAAAGATGCTATCCGGAATTCCTGCATACAG 1450
      |||
QY 1451 AAAAGCTCCACCATTTCTTTGATGCTTTAAAGTCTCAAGTCTCTATATATGAA 1510
      |||
Db 93741 AAAAGCTCCACCATTTCTTTGATGCTTTAAAGTCTCAAGTCTCTATATATGAA 1510
      |||
QY 1511 ACAGACAGCTCTGTCAGCTCTGCTGCTGATGCTGGAATGATGATGCTTACAGA 1570
      |||
Db 93681 ACAGACAGCTCTGTCAGCTCTGCTGCTGATGCTGGAATGATGATGCTTACAGA 1570
      |||
QY 1571 AAGCATTTTCTTTCTTTAACCCTTAAGCTCTATATTAAGACAGACAGATTCGA 1630
      |||
Db 93621 AAGCATTTTCTTTCTTTAACCCTTAAGCTCTATATTAAGACAGACAGATTCGA 1630
      |||
QY 1631 CATTTTATACATGAGATCTTCTTGTGCTGATACACAGATTCGATCCCTTAA 1690
      |||
Db 93561 CATTTTATACATGAGATCTTCTTGTGCTGATACACAGATTCGATCCCTTAA 1690
      |||
QY 1691 AAGAAGTTTATGTCCTGCTGCTGCTAAATTAATCTAATTCAGATCTTTTGTAGA 1750
      |||
Db 93502 AAGAAGTTTATGTCCTGCTGCTGCTAAATTAATCTAATTCAGATCTTTTGTAGA 1750
      |||
QY 1751 TGACAGATTTGTTGAGGACACATATGAGATCTAGATTTGAGTGAATGAGCAAGAA 1810
      |||
Db 93442 TGACAGATTTGTTGAGGACACATATGAGATCTAGATTTGAGTGAATGAGCAAGAA 1810
      |||
QY 1811 GGGCCATCTCCATGAGATGATTAAGTGAACCAACTAGTCTCGAATTCACAGAA 1870
      |||
Db 93382 GGGCCATCTCCATGAGATGATTAAGTGAACCAACTAGTCTCGAATTCACAGAA 1870
      |||
QY 1871 GGAGGGAATAGACTGAGGAACTGTGACATAGACTTGAAGACCAAGACTTTGAAT 1930
      |||
Db 93322 GGAGGGAATAGACTGAGGAACTGTGACATAGACTTGAAGACCAAGACTTTGAAT 1930
      |||
QY 1931 TGGGAGCTGCTCAGTGTGATGATTAATCACTGCTGCTTCTTATGATTACAAATCTA 1990
      |||
Db 93262 TGGGAGCTGCTCAGTGTGATGATTAATCACTGCTGCTTCTTATGATTACAAATCTA 1990
      |||
QY 1991 TATTTTATGAGTTTAATAAGAAAAAATTACAGA 2030
      |||
Db 93202 TATTTTATGAGTTTAATAAGAAAAAATTACAGA 93163
      |||

```

## JOURNAL

## COMMENT

Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 On Aug 21, 2000 this sequence version replaced gi:9213379.  
 ----- Genome Center  
 Center: Sanger Centre  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquerry@sanger.ac.uk  
 ----- Project Information  
 Center project name: BA473P22  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Consensus quality: 189476 bases at least Q40  
 Consensus quality: 195311 bases at least Q30  
 Consensus quality: 195510 bases at least Q20  
 Insert size: 198363; sum-of-coverage  
 Insert coverage: 201159; 6.2% error; agarose-fp  
 Quality coverage: 3.87x in Q20 bases; sum-of-coverage  
 coverage: 3.98x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 16 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 2278: contig of 2278 bp in length
2279 2378: gap of 100 bp
2379 2274: contig of 2035 bp in length
2274 2284: gap of 100 bp
2284 45183: contig of 22340 bp in length
45183 45283: gap of 100 bp
45283 49581: contig of 4298 bp in length
49581 49681: gap of 100 bp
49681 58258: contig of 8577 bp in length
58258 58358: gap of 100 bp
58358 64082: contig of 5724 bp in length
64082 64182: gap of 100 bp
64182 72764: contig of 8582 bp in length
72764 72864: gap of 100 bp
72864 86552: contig of 13688 bp in length
86552 86652: gap of 100 bp
86652 93481: contig of 6829 bp in length
93481 93581: gap of 100 bp
93581 101614: contig of 8033 bp in length
101614 101714: gap of 100 bp
101714 107836: contig of 6122 bp in length
107836 107936: gap of 100 bp
107936 144546: contig of 36610 bp in length
144546 144646: gap of 100 bp
144646 155892: contig of 11246 bp in length
155892 155992: gap of 100 bp
155992 167875: contig of 11883 bp in length
167875 167975: gap of 100 bp
167975 183411: contig of 15436 bp in length
183411 183511: gap of 100 bp
183511 199863: contig of 16352 bp in length.

```

## FEATURES

## source

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP11-473P22"  
 /clone\_1lb="RP11-11.2"

## misc-feature

1. 2278  
 /note="assembly\_fragment:00765  
 fragment\_chain:1  
 clone\_end:T7

1149 2470: contig of 1322 bp in length  
2471 2570: gap of 100 bp  
2571 3643: contig of 1073 bp in length  
3644 3743: gap of 100 bp  
3744 4980: contig of 1237 bp in length  
4981 5080: gap of 100 bp  
5081 6352: contig of 1272 bp in length  
6353 6452: gap of 100 bp  
6453 7549: contig of 1097 bp in length  
7550 7649: gap of 100 bp  
7650 8977: contig of 1328 bp in length  
8978 9077: gap of 100 bp  
9078 10103: contig of 1026 bp in length  
10104 10203: gap of 100 bp  
10204 11362: contig of 1159 bp in length  
11363 11462: gap of 100 bp  
11463 12821: contig of 1359 bp in length  
12822 12921: gap of 100 bp  
12922 14007: contig of 1086 bp in length  
14008 14107: gap of 100 bp  
14108 15283: contig of 1176 bp in length  
15284 15383: gap of 100 bp  
15384 16399: contig of 1016 bp in length  
16400 16499: gap of 100 bp  
16500 17804: contig of 1305 bp in length  
17805 17904: gap of 100 bp  
17905 19176: contig of 1272 bp in length  
19177 19276: gap of 100 bp  
19277 20384: contig of 1108 bp in length  
20385 20484: gap of 100 bp  
20485 21689: contig of 1205 bp in length  
21690 21789: gap of 100 bp  
21790 23191: contig of 1402 bp in length  
23192 23291: gap of 100 bp  
23292 24513: contig of 1222 bp in length  
24514 24613: gap of 100 bp  
24614 26135: contig of 1522 bp in length  
26136 26235: gap of 100 bp  
26236 27373: contig of 1138 bp in length  
27374 27473: gap of 100 bp  
27474 28723: contig of 1250 bp in length  
28724 28823: gap of 100 bp  
28824 30217: contig of 1394 bp in length  
30218 30317: gap of 100 bp  
30318 32056: contig of 1739 bp in length  
32057 32156: gap of 100 bp  
32157 33818: contig of 1662 bp in length  
33819 33918: gap of 100 bp  
33919 35156: contig of 1238 bp in length  
35157 35256: gap of 100 bp  
35257 36712: contig of 1456 bp in length  
36713 36812: gap of 100 bp  
36813 38379: contig of 1567 bp in length  
38380 38479: gap of 100 bp  
38480 40190: contig of 1711 bp in length  
40191 40290: gap of 100 bp  
40291 41780: contig of 1490 bp in length  
41781 41880: gap of 100 bp  
41881 44053: contig of 2173 bp in length  
44054 44153: gap of 100 bp  
44154 45474: contig of 1321 bp in length  
45475 45574: gap of 100 bp  
45575 46821: contig of 1247 bp in length  
46822 46921: gap of 100 bp  
46922 47302: contig of 381 bp in length  
47303 47402: gap of 100 bp  
47403 49323: contig of 1921 bp in length  
49324 49423: gap of 100 bp  
49424 51140: contig of 1717 bp in length  
51141 51240: gap of 100 bp  
51241 53099: contig of 1859 bp in length  
53100 53199: gap of 100 bp  
53200 54631: contig of 1432 bp in length

54632 54731: gap of 100 bp  
54732 55751: contig of 1020 bp in length  
55752 55851: gap of 100 bp  
55852 57470: contig of 1619 bp in length  
57471 57570: gap of 100 bp  
57571 60182: contig of 2612 bp in length  
60183 60282: gap of 100 bp  
60283 63065: contig of 2783 bp in length  
63066 63165: gap of 100 bp  
63166 64443: contig of 1278 bp in length  
64444 64543: gap of 100 bp  
64544 66043: contig of 1500 bp in length  
66044 66143: gap of 100 bp  
66144 68623: contig of 2480 bp in length  
68624 68723: gap of 100 bp  
68724 70122: contig of 1399 bp in length  
70123 70222: gap of 100 bp  
70223 72407: contig of 2185 bp in length  
72408 72507: gap of 100 bp  
72508 74635: contig of 2128 bp in length  
74636 74735: gap of 100 bp  
74736 77554: contig of 2819 bp in length  
77555 77654: gap of 100 bp  
77655 80277: contig of 2623 bp in length  
80278 80377: gap of 100 bp  
80378 81798: contig of 1421 bp in length  
81799 81898: gap of 100 bp  
81899 84126: contig of 2228 bp in length  
84127 84226: gap of 100 bp  
84227 86794: contig of 2568 bp in length  
86795 86894: gap of 100 bp  
86895 90311: contig of 3417 bp in length  
90312 90411: gap of 100 bp  
90412 92255: contig of 1844 bp in length  
92256 92355: gap of 100 bp  
92356 94570: contig of 2215 bp in length  
94571 94670: gap of 100 bp  
94671 97163: contig of 2493 bp in length  
97164 97263: gap of 100 bp  
97264 99724: contig of 2461 bp in length  
99725 99824: gap of 100 bp  
99825 103151: contig of 3327 bp in length  
103152 103251: gap of 100 bp  
103252 106370: contig of 3119 bp in length  
106371 106470: gap of 100 bp  
106471 108886: contig of 2418 bp in length  
108887 108988: gap of 100 bp  
108889 111514: contig of 2526 bp in length  
111515 111614: gap of 100 bp  
111615 114360: contig of 2746 bp in length  
114361 114460: gap of 100 bp  
114461 118269: contig of 3809 bp in length  
118270 118369: gap of 100 bp  
118370 122359: contig of 3990 bp in length  
122360 122459: gap of 100 bp  
122460 125669: contig of 3210 bp in length  
125670 125769: gap of 100 bp  
125770 129306: contig of 3537 bp in length  
129307 129406: gap of 100 bp  
129407 132984: contig of 3578 bp in length  
132985 133084: gap of 100 bp  
133085 137360: contig of 4276 bp in length  
137361 137460: gap of 100 bp  
137461 140476: contig of 3016 bp in length  
140477 140576: gap of 100 bp

Query Match 41.7% Score 850.4; DB 2; Length 200484;  
Best Local Similarity 98.6% Pred. No. 6; E-211;  
Matches 868; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
OY 1151 GATCGCATGCGAGCGACATATAATCCAAATCCACCTGGAACACGAGAGAGAGA 1210  
DB 94041 GATTTCGCTATACAGCGCCATATAATCCAAATCCCACTTGAACCAACTGAGAGAGAGA 93982



<http://bacpac.med.buffalo.edu/>  
 VECTOR: pcYpAC2  
 This sequence is the entire insert of clone RPI-118J21. The true  
 right end of clone RPI-11/L23 is at 38518 in this sequence. This  
 sequence was generated from part of bacterial clone contigs of  
 human chromosome 1, constructed by the Sanger Centre Chromosome 1  
 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr1>.

## FEATURES

SOURCE  
 1. 166496  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="p34.1-35.3"  
 /clone="RPI-118J21"  
 /clone\_11b="RPI-1"  
 complement(830..29381)  
 gene  
 /gene="BMP8"  
 complement(join(<830..1098,3622..3732,4242..4321,  
 5153..5347,14838..14986,15379..15568,28682..29381))  
 /gene="BMP8"  
 /product="DJ118J21.1 (bone morphogenetic protein 8  
 (osteogenic protein 2))"  
 /note="match: CDNAS: Em:S77477 Em:X51801 Em:M97017  
 Em:U39545 Em:J04566 Em:X80992 Em:M60316 Em:M60315  
 Em:M97016 Em:X56906 Em:AF100787 Em:U40034"  
 /evidence="not\_experimental"  
 complement(join(949..1098,3622..3732,4242..4321,  
 5153..5347,14838..14986,15379..15568,28682..29015))  
 /gene="BMP8"  
 /note="match: proteins: Sw:P22004 Sw:P20722 Sw:P22003  
 Sw:P18075 Tr:O9YGH7 Sw:P23359 Sw:P55105 Sw:P49003  
 Sw:P34820 Sw:P34821"  
 /codon\_start=1  
 /evidence="not\_experimental"  
 /product="DJ118J21.1 (bone morphogenetic protein 8  
 (osteogenic protein 2))"  
 /protein\_id="CAB75681.1"  
 /db\_xref="gi:7018292"  
 /translation="MTALPGLPLMLGLALCALGGGPELRPPGCPORRLCARERRDV  
 QRETLAVLGLGRPRAPAPASRLPAPASPLFMDLYAMAGDDDEDAPAEERRIGRA  
 DLVNSFVNMVERDALGHOEPMKREFEDLQIAPAGEVTAAEFRIVKPSHILNRT  
 LHVNSFVQVQSNRESDFLFDLOTLAGDEGCVLDVTAASDCMLKRRKDLGLRL  
 YVEREDGSDVPGLAGLIGRAPRSQGFVFPASPIRTPRAPVPLRRDROKKS  
 NELPQANRIPETPDVYRSHRGVCRRLHLYVSFDCLMDLVIAPOGTSVYCEGEC  
 SFPLDSKNATNHA1LOSVLVHMPDAVPAKCACTKATSVLYDSSNNVTLKKHR  
 NMVYKACGCH"  
 1655..1748  
 repeat\_region  
 /note="L2 repeat: matches 1691..1787 of consensus"  
 1849..2281  
 repeat\_region  
 /note="L2 repeat: matches 2002..2438 of consensus"  
 2764..2868  
 repeat\_region  
 /note="FLAM\_C repeat: matches 1..133 of consensus"  
 complement(4214..4420)  
 /gene="BMP8"  
 /note="match: STS: Em:T63436"  
 complement(4252..4438)  
 /gene="BMP8"  
 /note="match: STS: Em:AA634039"  
 5972..6029  
 repeat\_region  
 /note="MIR repeat: matches 19..75 of consensus"  
 6293..6588  
 repeat\_region  
 /note="Aluub repeat: matches 1..307 of consensus"  
 7787..7873  
 repeat\_region  
 /note="MIR repeat: matches 57..146 of consensus"  
 complement(8369..8876)  
 /gene="BMP8"  
 /note="match: GSS: Em:B51067"  
 9109..9404  
 repeat\_region  
 /note="Aluub repeat: matches 1..296 of consensus"  
 9427..9746  
 repeat\_region  
 /note="Aluub repeat: matches 1..312 of consensus"  
 10075..10242  
 misc\_feature

misc\_feature  
 /note="match: STS: Em:Z41229"  
 10626..12048  
 /note="Cpg Island"  
 /evidence="not\_experimental"  
 repeat\_region  
 12351..13402  
 /note="L2 repeat: matches 957..2141 of consensus"  
 repeat\_region  
 13474..13591  
 /note="Aluuo/FRAM repeat: matches 178..295 of consensus"  
 repeat\_region  
 13596..14185  
 /note="L1MC3 repeat: matches 7071..7691 of consensus"  
 repeat\_region  
 14186..14425  
 /note="L2 repeat: matches 2129..2417 of consensus"  
 repeat\_region  
 14457..14572  
 /note="MIR repeat: matches 27..147 of consensus"  
 repeat\_region  
 15137..15215  
 /note="MIR repeat: matches 47..137 of consensus"  
 repeat\_region  
 16285..16510  
 /note="L2 repeat: matches 2361..2586 of consensus"  
 repeat\_region  
 17313..17597  
 /note="L2 repeat: matches 2445..2745 of consensus"  
 repeat\_region  
 20320..20782  
 /note="MIR repeat: matches 3..541 of consensus"  
 repeat\_region  
 20785..20932  
 /note="MIR repeat: matches 22..175 of consensus"  
 repeat\_region  
 21140..21425  
 /note="Aluux repeat: matches 1..282 of consensus"  
 repeat\_region  
 21927..22073  
 /note="MIR repeat: matches 97..262 of consensus"  
 repeat\_region  
 22185..22408  
 /note="L2 repeat: matches 116..325 of consensus"  
 repeat\_region  
 22409..22704  
 /note="AluSg repeat: matches 5..306 of consensus"  
 repeat\_region  
 22705..22759  
 /note="L2 repeat: matches 60..116 of consensus"  
 repeat\_region  
 22829..22898  
 /note="MIR repeat: matches 5..71 of consensus"  
 repeat\_region  
 22952..22989  
 /note="MIR repeat: matches 227..261 of consensus"  
 repeat\_region  
 22990..23298  
 /note="AluSg repeat: matches 1..308 of consensus"  
 repeat\_region  
 23299..23381  
 /note="MIR repeat: matches 132..227 of consensus"  
 repeat\_region  
 23401..23576  
 /note="Aluux repeat: matches 51..223 of consensus"  
 repeat\_region  
 23600..23742  
 /note="MIR repeat: matches 8..153 of consensus"  
 repeat\_region  
 24157..24451  
 /note="Aluuo repeat: matches 2..294 of consensus"  
 repeat\_region  
 24618..24684  
 /note="MIR repeat: matches 80..147 of consensus"  
 repeat\_region  
 24753..24891  
 /note="MIR repeat: matches 50..199 of consensus"  
 repeat\_region  
 25018..25159  
 /note="L2 repeat: matches 2606..2750 of consensus"  
 repeat\_region  
 26215..26275  
 /note="AluSg repeat: matches 143..195 of consensus"  
 repeat\_region  
 26604..26643  
 /note="MIR repeat: matches 80..119 of consensus"  
 repeat\_region  
 26930..27021  
 /note="MIR repeat: matches 81..179 of consensus"  
 repeat\_region  
 27580..27733  
 /note="MIR repeat: matches 37..197 of consensus"  
 repeat\_region  
 28059..28151  
 /note="L2 repeat: matches 2598..2703 of consensus"  
 repeat\_region  
 28542..30030  
 /note="Cpg Island"  
 /evidence="not\_experimental"  
 repeat\_region  
 29451..29584  
 /note="67 copies 2 mer gg 598 conserved"  
 repeat\_region  
 30126..30262  
 /note="L2 repeat: matches 2350..2475 of consensus"  
 repeat\_region  
 30283..30460  
 /note="MIR repeat: matches 60..241 of consensus"

QY	569	AGGAGCTGTGCAAGTTTGTGAAAGAACAGAAATCTGCATAGTGAATTTCTCCATGCTAC	628
Db	541	AGGAGCTGTGCAAGTTTGTGAAAGAACAGAAATCTCTCACAGTGAATTTCTTCATCTGAC	600
QY	629	CATACGGAAGAGTGGTGGTCCCTTGGAGGTCCTGTGAAATTTCTTAACCTTGCATC	688
Db	601	CACGACAGGAAGAGCGGTGGTCTCTTTGGAGGCCCTCTCAGGTCTCCTAACCCATGATC	660
QY	689	CTTTGGCTTATGCTGACCAGGCAGTTCTAGATGAGCCGTTGGATPAAGGGGTGATGAC	748
Db	661	CTCTGGCTTCATCTGTGACAGGCAGTTTCTTAGATGAGCCGTTGGATPAAGAGGTGATGAC	720
QY	749	ATGCTTCTCTGGGCTCTGGAGGAACATAGAGATTTTACAGACGGTATPATATAGAG	808
Db	721	ATGCTTCTCTCGGACTCTTGGAGAGCTGAGAGGTTTTCACAGACGTTATPATATTAG	780
QY	809	AATGTTTGGAAAAATAGCCAGGACTATCAACATGGATTTCTCAATCAATTGGCTTACAG	868
Db	781	AACATTTCAAAAAATATAGCCAGGACTATCAACAGGATTTCTCACTCAATTGGCTTACAG	840
QY	869	GAAATTTACAGATGCTGATCACTAGGAGGAATGACACTGAGACATGATTAACCACTT	928
Db	841	GAAATTTACAGATGCTGATCACTAGGAGGAATGACACACAGAGACTGATTAACCACTT	900
QY	929	CTAAGGAAG-----	938
Db	901	CTAAGGAAGAGTATGAGGCTGTGAAACAAGTAACTAAGATATATCCGGAAACAGAAC	960
QY	939	-----GACCTGGTCCCATTTGTCGCCCTCTGTATGAGC	970
Db	961	CGATGGGTTAAAAACGCTTTTGGACAGACCTGGGCCCAAGTGTCCCCCGGTATATGCG	1020
QY	971	TTAGAGGTATCTGATGTCTGGAAGTGGGAGAGCTGTGTTCTTGAACCTGCTCTTGAATC	1030
Db	1021	TTAGAGGTATCTGATGTGTTTCCAGGTGGGAGAGACTGTGTTCTGGAACCTGCTCTCAATTA	1080
QY	1031	GTCGAAAGTTTCATCCAGGGCCACAAGCCTACAGCCACTGCATTAAGATGCCATACAT	1090
Db	1081	GTCGAAAGTTTCATCCAGGGCTTCAAAAACCTACAGCCATCCAGTGAAGATGGCATACAT	1140
QY	1091	GAACTGTGGAACAAGACAACTTTACACTGCTGTGACCTCTGTGATGGAATCATCTATGG	1150
Db	1141	GAAAGTGTGAACAAGAAATATACCAATGTGACCTCTGTGACCGGATCTATCTATGG	1200
QY	1151	GATCGCAATTTGGGAGGCGACATTAATATCCAAATCCCACTTAACCACTGAGAAAAAGA	1210
Db	1201	GACCGGGAATGGGAGACACTTTAAATATCCAAATCTCACTTGACCAATTTGAAGAAAAGA	1260
QY	1211	AGAGAAATTGCATCAGTGTCTGTCAACACCATTAAGAAAGTCAAGTGTGTTCCCGACATAT	1270
Db	1261	AGAGAAATTGCATCAGCTGTGTACAGTGTGTGACGACAGGAAGTCAAGTATATCCCGAACAATG	1320
QY	1271	AACAAAGAACTTAAGGGAAGGATATCCCAAGGGCACAATATGATCAAGAGCTGAAATGCAAC	1330
Db	1321	GACCGGAACCAATCGAGGGGGAATCTCTGGGGCACACAAATCAAGAGCTGAAAGGCCAGT	1380
QY	1331	GTTTAAAGACATGTCCAGTGGGCTTTGGAAAAAGTGGTGGATCCAGTTCAGAGAGGAG	1390
Db	1381	GTTTGAAGACATGTCCATAGTGGCTTTTGCAAGA-GACGTGGGATTCAGAGCCACAGAGGAG	1430
QY	1391	GGGTATGTTGTCTCCCA-GTCTGGGCAAGAGAGTCTATGGGGAATTTCTGCAATAGCA	1440
Db	1440	GGGAGTGTGCGCTCTCCACAGCGCTGGAATGAGAAATCTGTGGC-----	1480
QY	1450	GAAAGGCTCCCACTTTTCTTTTATGTGTGTTTAAAGTCTCAGCTTCTCTATATATAGA	1500
Db	1482	AGAGAGGCCCACTATCTTCTTCAATCTG-----TGCTGTGCTGTGCAGTGGA	1520
QY	1510	AACGACAGGTTTGTACAGTCCTTGTGTGCTGATGCTGTGCTGGAATATGATATGATTCAG	1560
Db	1530	AACGACAGGCTT-TGAGCTCTTGTGTGCTCTTGTGTGTGTGATATGATATGATATGATTCAG	1580

QY	1570	AAACATTTTTCCTTTCCTTACCTTAAGCTTAAGTCTATTTATTAAGACGACAGATCC	1629
Db	1589	GTGGGA---TTTTCCTTTCAACCTTAAGCTTTATTTTGAACGACGACAGATGCG	1645
QY	1630	ACATTTTATATACATAGAGATCTCTTTGTGTGAATACCAAGATTGACTGCATCCCTT	1689
Db	1646	ACATTTCTACTTACGTAGATCTTTTATGTGTGAATACCAAGATTCACTGATC-CC	1704
QY	1690	AAACATGTTTATATCCCTGCCTCGCTAAATATATCTAATTCACATGCTTTGTAG	1749
Db	1705	AAAGAGCTTTCTTGCTCCCTGGCGCTGGCTAAATAATAGCTCGTTTCAAGATGTTTGTAG	1764
QY	1750	ATGACTGAAGTATTTGTGTAGCCACATATTGGAGTTCTAGATTGTGAGTAATGGCAGAA	1809
Db	1765	ATGACTGAAGTATTTGTAGCGCAGCTCAGAGATCTGAGCTCGACAGAACGGCAGAAAGGAG	1824
QY	1810	AGGGCATCTCCATTTGAGATGATTAAATGAACCAACACTAGTCTTCGGAAATTCACAGGA	1869
Db	1825	TAG-----TCCAGAGATATGATTAAAGGAACCTCCACGCTCTGTGAT---CTACAG	1874
QY	1870	AGGAGGGAATGAGACTGAGAAAGCTGTACATATAGCACTTGAAGAC-CAAGACTTTGAAA	1928
Db	1875	AAGAGGGGTGACGGCTGAGGTGTTGTATACCTCGATCTTTGAAGATTAAGAACAAGAGT	1934
QY	1929	TTTGGAGCTGCTCATGTGTGAGTTATTAATCACTGCTGCTTTCTATTGAGTTACAAATC	1988
Db	1935	CTGCTGACCTGCTCTGTGCATGTGGCTGTCTTTCATCTGAGGACAGAAATCATATAGT	1994
QY	1989	TATATTTTATTTGAAGTTTAATATAAGAAAAAATTACAGAAAAAATAA 2039	
Db	1995	TCCTTTGAATAATATAAAGATTTGTAAACGCCCAAAAAAAAAAAAAA 2045	

## RESULT 9

[illegible]

**COMMENT**

On Nov 21, 1999 this sequence version replaced g1:60022299. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences without a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RPL-118J21 is from the library RPL-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see

QY	669	CTTGGCCTCAATGCGAGCAGGAGCTTCTACATGAGCGCTTGAAATGAAGGGTGGATAC	748
Db	661	CTTTGGCTTCAATGCTGACGACGAGCAGTTCTAATATGAGCGCTTGAAATGAAGGGTGGATAC <td>720</td>	720
QY	749	ATGCTTGCCTGCTGGGCTCTTTGAGAGACATTAAGAGATTTTTCACAGACGCTATATACAGAAG	808
Db	721	ATGCTTGCCTGCTGGGCTCTTTGAGAGACATTAAGAGATTTTTCACAGACGCTATATACAGAAG	780
QY	809	AATGTTTCGGAATAATAGCCAGGACTATACACATGGTATCTTCCAAATCAATTGGCTTCAAG	868
Db	781	AATGTTTCGGAATAATAGCCAGGACTATACACATGGTATCTTCCAAATCAATTGGCTTCAAG	840
QY	869	GAATTTTCACGAGTACCTATACATGAGGAGAAATGACACATGAGAGACTAGTATACCAGCTT	928
Db	841	GAATTTTCACGAGTACCTATACATGAGGAGAAATGACACATGAGAGACTAGTATACCAGCTT	900
QY	929	CTTAAGAAAGAACCGGATGCGCCATGTCCTCCCTGTCCTATGGCTTATGAGTATGATGTC	988
Db	901	CTTAAGAAAGAACCGGATGCGCCATGTCCTCCCTGTCCTATGGCTTATGAGTATGATGTC	960
QY	989	TGCAAGTGGGAGGAGTCTGTTCTTGAACCTGCTCTTGAATCTGTCAGAAAGTTTCAATCCAG	1048
Db	961	TGCAAGTGGGAGGAGTCTGTTCTTGAACCTGCTCTTGAATCTGTCAGAAAGTTTCAATCCAG	1020
QY	1049	GGCCACAGAGCCTTACAGCCACTCCCAATAAAGATGCCATCAATGAAGCTGAGAACACAGAGA	1108
Db	1021	GGCCACAGAGCCTTACAGCCACTCCCAATAAAGATGCCATCAATGAAGCTGAGAACACAGAGA	1080
QY	1109	AGTTATCAACCTGTGACACTCTGTGATGTGAATCATCATTTGGGGATGCGGAATGGGCACGC	1168
Db	1081	AGTTATCAACCTGTGACACTCTGTGATGTGAATCATCATTTGGGGATGCGGAATGGGCACGC	1140
QY	1169	CACATTAATCAATCAATCCCACTTGAACCACTGGAAGAAAGAAAGAAATGAGTACTCAGAT	1228
Db	1141	CACATTAATCAATCAATCCCACTTGAACCACTGGAAGAAAGAAAGAAATGAGTACTCAGAT	1200
QY	1229	GCTGTCAACACCATTAGAAAGTCAAGATGTTTCCCGACACTATACAAAGAACTTAAGAGG	1288
Db	1201	GCTGTCAACACCATTAGAAAGTCAAGATGTTTCCCGACACTATACAAAGAACTTAAGAGAG	1260
QY	1289	AAGGATCCCGCAGGCGAGATGATCAAGAGCTGAATGCAATGACCGTTTAA	1336
Db	1261	AAGGATCCCGCAGGCGAGATGATCAAGAGCTGAATGCAATGACCGTTTAA	1308
RESULT 8			
LOCUS	BC019812		
DEFINITION	Mus musculus, RIKEN cDNA 2310075614 gene, clone MGC:30541		
ACCESSION	IMAGE:5042856, mRNA, complete cds.		
VERSION	BC019812		
KEYWORDS	BC019812.1 GI:18044185		
SOURCE	MGC.		
ORGANISM	house mouse.		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 2045)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-DEC-2001) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.ncl.nih.gov">http://mgc.ncl.nih.gov</a>		
COMMENT	Contact: MGC help desk		
	Email: <a href="mailto:cgabs-remail.nih.gov">cgabs-remail.nih.gov</a>		
	Tissue Procurement: Jeffrey E. Green, M.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Baylor College of Medicine Human Genome		
	Sequencing Center		
	Center code: BCM-HGSC		

```

|||||
Db 1109 AGAATGATTCAGACGCTTAATGACGCTTTAGAGACATGCCAGTGGCTTTGGAAGG 1168
QY 1365 TGGTGGGGATTCAGATTCAGAGAGGAGGATGTTGTTGTCGCCAGTGGCAAGAGAGT 1424
Db 1169 TGGTGGGGATTCAGATTCAGAGAGGAGGATGTTGTTGTCGCCAGTGGCAAGAGAGT 1228
QY 1425 GCTATGCGGAATTCCTGTCATAGCAGAAAAGCTCCACCATTTCTTTTGATGATGTTT 1484
Db 1229 GCTATGCGGAATTCCTGTCATAGCAGAAAAGCTCCACCATTTCTTTTGATGATGTTT 1288
QY 1485 AAATCTCAGGTTCTCTATATAGAACAGAGGCTTGTGACGCTCCCTGTGGTGGAT 1544
Db 1289 AAATCTCAGGTTCTCTATATAGAACAGAGGCTTGTGACGCTCCCTGTGGTGGAT 1348
QY 1545 GTGCTCGAATGATGATGATTCAGAAAAGCATTTTCTTTTCTTGAACCTTAAAGGTTT 1604
Db 1349 GTGCTCGAATGATGATGATTCAGAAAAGCATTTTCTTTTCTTGAACCTTAAAGGTTT 1408
QY 1605 TATATTTAAAGACGACACATTCACATTTTATACATGAGATCTTCTTTGTGTGA 1664
Db 1409 TATATTTAAAGACGACACATTCACATTTTATACATGAGATCTTCTTTGTGTGA 1468
QY 1665 TACCAGATTCAGTGCATCCCTTTAAAGAGTTTATGTCGCCGACCTGGCTAAATTT 1724
Db 1469 TACCAGATTCAGTGCATCCCTTT-AAAGAGTTTATGTCGCCGACCTGGCTAAATTT 1527
QY 1725 ATCTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATGATGAT 1784
Db 1528 ATCTAATTTCCAGATGCTTTTGTAGATGATGATGATGATGATGATGATGATGATGAT 1587
QY 1785 TCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
Db 1588 TCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1647
QY 1845 ACTAGTTCTCGAATTCCTACAGAGAGAGGAGTTCAGACGAGAGAGCTGTGACATAG 1904
Db 1848 ACTAGTTCTCGAATTCCTACAGAGAGAGGAGTTCAGACGAGAGAGCTGTGACATAG 1707
QY 1905 ACTGGAAGACCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGATGATGATGATGATGAT 1964
Db 1708 ACTGGAAGACCAAGACTTTGAAATTTGCGAGCTGCTCATGCTGATGATGATGATGATGAT 1767
QY 1965 TGTCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2024
Db 1768 TGTCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1827
QY 2025 ACAGAAAAAAGAAAAA 2041
Db 1828 ACAGAAAAAAGAAAAA 1844

RESULT 7
LOCUS AY052768 1308 bp mRNA linear PRI 17-OCT-2001
DEFINITION Homo sapiens tRNA isopentenyl transferase mRNA, partial cds.
ACCESSION AY052768
VERSION AY052768.1 GI:16209578
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Lemieux,J., Lakowski,B., Webb,A., Meng,Y., Udech,A., Bussiere,F.,
Barnes,T. and Hekimi,S.
Regulation of physiological rates in Caenorhabditis elegans by a
tRNA-modifying enzyme in the mitochondria
Genetics 159 (1), 147-157 (2001)
JOURNAL MEDLINE 2144833
PUBMED 11560893
REFERENCE 2 (bases 1 to 1308)
AUTHORS Lemieux,J., Barnes,T. and Hekimi,S.

```

```

TITLE Direct Submission
JOURNAL Submitted (28-AUG-2001) Biology, McGill University, 1205 Dr.
Pentfield Avenue, Montreal, QC H3A 1B1, Canada
FEATURES
source 1..1308
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="c-2ec05"
<1..1308
/note="hgr01"
/codon_start=1
/product="tRNA isopentenyl transferase"
/protein_id="AA14107.1"
/db_xref="GI:16209579"
/translation="AARVPVGSGLRGLRPLVILGATGKSTIALQLQRLGE
IVSADSMQVEGLDITNNKSAQDORICRHHMISFVPLVNTYVDFRNRATLIED
IFARDKPIVVGKNYIESLILKLVNTPEQETGKVIDRKELEKEDGLVLRKL
SOYDEPAALPHDKRVARSLQVFEETGSEFLRPHQTEGGGLGGLPKFSNP
CIIMLHADOAVLDRDKRVDMLAAGLELRLPHFRHRYNOKNYSNSODYOHGIFPS
IGREHEHEYLITBEKCTLENSNOLKRGPRPIYPVYGLEYSVSKESYLEPALKI
VOSTGCHKPTATPKMPYDEAKNSYHLCDCDRIIGREMAHAIKSKSHLQK
KRRRLDSDAVNTIESQSVSDYKNREPKESPQNDDELKCSV"
BASE COUNT 378 a 302 c 329 g 299 t
ORIGIN
Query Match 64.0%; Score 1306.4; DB 9; Length 1308;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 29 GCAGAGCAGTTCCTGTCGTCGAGTGGGCTCAGAGGAGCTGCAAGGAGACCTTACTGTA 88
Db 1 GCAGAGCAGTTCCTGTCGTCGAGTGGGCTCAGAGGAGCTGCAAGGAGACCTTACTGTA 60
QY 89 GTGATTCGAGGACACGCGGACCGCAATCCAGCTGGCGTTGCACCTAGCCAGCGG 148
Db 61 GTGATTCGAGGACACGCGGACCGCAATCCAGCTGGCGTTGCACCTAGCCAGCGG 120
QY 149 CTCGGCGGTGAGTGTGTAGGCTACTCCATGCAGTCTATGAAAGGCTTAGACATCATC 208
Db 121 CTCGGCGGTGAGTGTGTAGGCTACTCCATGCAGTCTATGAAAGGCTTAGACATCATC 180
QY 209 ACCAACAAGGTTTCCGCCAAGAGAGAGAAATCTCCGCAACCATATGATGATGATGATGAT 268
Db 181 ACCAACAAGGTTTCCGCCAAGAGAGAGAAATCTCCGCAACCATATGATGATGATGATGAT 240
QY 269 GATCCTCTTGACCAATTAACACAGTGTGACATTCAGAAATAGACAACTGCTGATTT 328
Db 241 GATCCTCTTGACCAATTAACACAGTGTGACATTCAGAAATAGACAACTGCTGATTT 300
QY 329 GAGATATATTTGCCCGAGACAAATTCCTATTTGTTGGAGAGAACCAATTTATACATT 388
Db 301 GAGATATATTTGCCCGAGACAAATTCCTATTTGTTGGAGAGAACCAATTTATACATT 360
QY 389 GAATCTGCTGTGAAAGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACTGAGAA 448
Db 361 GAATCTGCTGTGAAAGTTCTTGTCAATACCAAGCCCGAGAGATGGGCACTGAGAA 420
QY 449 GTGATTCAGCAAAAGTGTGAGAGAGAGAGATGCTTGTACTTCACAAAGCCCTA 508
Db 421 GTGATTCAGCAAAAGTGTGAGAGAGAGATGCTTGTACTTCACAAAGCCCTA 480
QY 509 AGCCAGGTGACCCAGAAATGCTGCAAGCTGATCCATGACATGACAAACGCAAGTGGCC 568
Db 481 AGCCAGGTGACCCAGAAATGCTGCAAGCTGATCCATGACATGACAAACGCAAGTGGCC 540
QY 569 AGGAGCTTGCAAGTTTGAAGAAACAGAGATCTCTATAGTAATTTCTCCATGCTGTA 628
Db 541 AGGAGCTTGCAAGTTTGAAGAAACAGAGATCTCTATAGTAATTTCTCCATGCTGTA 600
QY 629 CATACGGAAGAGGAGTGTGCTGCTGAGAGTCTCTGAAAGTCTTAACCTTGATTC 688
Db 601 CATACGGAAGAGGAGTGTGCTGCTGAGAGTCTCTGAAAGTCTTAACCTTGATTC 660

```



Db 1560 GAATTCACGAGAGAGAGATCAGACTGAGAGAGCTGATCAGACTGAGAGACC 1619  
 QY 1916 AAGACCTTGAATTTGACAGCTGCTCATGTGTGATTTATTCATCTGCTTCTTAT 1975  
 Db 1620 AAGAGCTTGAATTTGACAGCTGCTCATGTGTGATTTATTCATCTGCTTCTTAT 1679  
 QY 1976 TGAGTTACAATCTATATTTTATTTGAGCTTTAAATTAAGAAAAATTTTACAGAAAAA 2035  
 Db 1680 TGAGTTACAATCTATATTTTATTTGAGCTTTAAATTAAGAAAAATTTTAAAAA 1739  
 QY 2036 AAAAAA 2041  
 Db 1740 AAAAAA 1745

RESULT 6  
 BC010741 1844 bp mRNA linear PRI 17-JUL-2001  
 LOCUS BC010741  
 DEFINITION Homo sapiens, similar to tRNA isopentenylpyrophosphate transferase,  
 clone MGC:17002 IMAGE:3905836, mRNA, complete cds.  
 ACCESSION BC010741  
 VERSION BC010741.1 GI:14789610  
 KEYWORDS MGC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 1844)  
 Strausberg, R.  
 Direct Submission  
 Submitted (12-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: cgabbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: http://www-shgc.stanford.edu  
 Contact: (Dickson, Mark) mdickpax1.stanford.edu  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

REMARK  
 COMMENT  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
 Series: IRAC Plate: 14 Row: P Column: 15  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 7019914.  
 Location/Qualifiers  
 1..1844  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="MGC:17002 IMAGE:3905836"  
 /tissue\_type="uterus, leiomyosarcoma"  
 /clone\_id="NIR\_MGC\_71"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6"  
 166..1140  
 /codon\_start=1  
 /product="similar to tRNA isopentenylpyrophosphate  
 transferase"  
 /protein\_id="AAH0741.1"  
 /db\_xref="GI:14789611"  
 /translation="MGTEKVIDRKVELEKEDGLVHKRLSDVPEMAAKLPHDNRKY  
 ASLSQFEETGSHSEFLRHQHTBEGGGLPLFNSNCILMLADODELDRVD  
 DMLAGLLELEDFHFRYRNOKNSQNSODYSGISGFKFHEYLITKEGCTLET  
 NDLKKGIEALKQVTKRYARKKRNKRYKRFSLRPPIYPVYGEVSDSKSESVLE  
 PALEIVQSFIOGHKPTATPIKMPYNEAKERKRYHLCDLCRIITIGDREWAHAKSKSH

## CDS

166..1140  
 /codon\_start=1  
 /product="similar to tRNA isopentenylpyrophosphate  
 transferase"  
 /protein\_id="AAH0741.1"  
 /db\_xref="GI:14789611"  
 /translation="MGTEKVIDRKVELEKEDGLVHKRLSDVPEMAAKLPHDNRKY  
 ASLSQFEETGSHSEFLRHQHTBEGGGLPLFNSNCILMLADODELDRVD  
 DMLAGLLELEDFHFRYRNOKNSQNSODYSGISGFKFHEYLITKEGCTLET  
 NDLKKGIEALKQVTKRYARKKRNKRYKRFSLRPPIYPVYGEVSDSKSESVLE  
 PALEIVQSFIOGHKPTATPIKMPYNEAKERKRYHLCDLCRIITIGDREWAHAKSKSH

BASE COUNT 545 a 374 c 454 g 471 t  
 ORIGIN

Query Match 73.0%; Score 1489; DB 9; Length 1844;  
 Best Local Similarity 94.7%; Pred. No. 0;  
 Matches 1607; Conservative 0; Mismatches 5; Indels 85; Gaps 3;

QY 423 ACCCCAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAAGCTTGAAGAGAG 482  
 Db 155 ACCCCAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGAAGAGAG 214  
 QY 483 ATGCTCTTACTTCACAAAGCCCTAAGCCAGGTGACCCAAATATGCTGCAAGCTGC 542  
 Db 215 ATGCTCTTACTTCACAAAGCCCTAAGCCAGGTGACCCAAATATGCTGCAAGCTGC 274  
 QY 543 ATCCACATGACAAAGCAAGTGGGCAAGGCTGCAAGTTTGAAGAAACAGGAACT 602  
 Db 275 ATCCACATGACAAAGCAAGTGGGCAAGGCTGCAAGTTTGAAGAAACAGGAACT 334  
 QY 603 CTCATAGTGAATTTCTCATCGTCAACATACGGAAGAGTGTGTCCTTGGAGGTC 662  
 Db 335 CTCATAGTGAATTTCTCATCGTCAACATACGGAAGAGTGTGTCCTTGGAGGTC 394  
 QY 663 CTCATAGTGTCTTAAACCTTCATCCTTGTGCTTCAAGTCAAGGCAAGTTTATGATG 722  
 Db 395 CTCATAGTGTCTTAAACCTTCATCCTTGTGCTTCAAGTCAAGGCAAGGCTTATGATG 448  
 QY 723 AGCGCTGATTAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 782  
 Db 449 AGCGCTGATTAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 508  
 QY 783 ATTTTACAGAGCTTAATCAAGAAATGTTTGGAAATAGCCAGAGCTATCAACATG 842  
 Db 509 ATTTTACAGAGCTTAATCAAGAAATGTTTGGAAATAGCCAGAGCTATCAACATG 568  
 QY 843 GATCTTCCATCAATTTGGCTTCAAGAAATTTACAGAGTACTGATGATGATGATGATGATG 902  
 Db 569 GATCTTCCATCAATTTGGCTTCAAGAAATTTACAGAGTACTGATGATGATGATGATGATG 628  
 QY 903 GCACATGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 938  
 Db 629 GCACATGAGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 688  
 QY 939 -----GACTG 944  
 Db 689 CTAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 748  
 QY 945 GTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1004  
 Db 749 GTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 808  
 QY 1005 CTTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064  
 Db 809 CTTGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 868  
 QY 1065 CCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124  
 Db 869 CCACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928  
 QY 1125 ACCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1184  
 Db 929 ACCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 988  
 QY 1185 CCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1244  
 Db 989 CCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1048  
 QY 1245 AAGTGAAGTGTGTTCCAGAGATTAACAAGAAAGTAAAGAGAGATGCCAGGCG 1304  
 Db 1049 AAGTGAAGTGTGTTCCAGAGATTAACAAGAAAGTAAAGAGAGATGCCAGGCG 1108  
 QY 1305 AGAATGATCAAGAGCTGAATGACGCTTTAAAGAGACATGTCAGTGCCTTTGGAAGG 1364

KEYWORDS oligo capling; fls (full insert sequence).

SOURCE Homo sapiens colon cDNA to mRNA, clone lib:COL clone:COL01371.

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Kawabata,A., Hiki,J.T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shihahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.

TITLE NEDO human cDNA sequencing project

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1749)

AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shihahara,T., Tanaka,T. and Nakamura,Y.

TITLE Direct Submission

JOURNAL Submitted (14-FEB-2002) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@leims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'-3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. 1749

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="COL01371"

/tissue\_type="colon"

/clone\_lib="COL"

/note="cloning vector: pME18FL3"

BASE COUNT 541 a 337 c 408 g 463 t

ORIGIN

Query Match 76.3%; Score 1556.8; DB 9; Length 1749;

Best Local Similarity 95.1%; Pred. No. 0;

Matches 1660; Conservative 0; Mismatches 7; Indels 79; Gaps 2;

374 ACCAATTATTTACCTGAATCTGCTGCTGGAAGTCTTGTCAATCCAGCCCGCAGAG 433

1 ACCAATTATTTACCTGAATCTGCTGCTGGAAGTCTTGTCAATCCAGCCCGCAGAG 60

434 ATGGGCACTGAGAAGATGATGACCGAAAAGTGGAGCTTGAAGAAGAGATGCTTGTGA 493

61 ATGGGCACTGAGAAGATGATGACCGAAAAGTGGAGCTTGAAGAAGAGATGCTTGTGA 120

494 CTTTCACAAACGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCATGAC 553

121 CTTTCACAAACGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATCCATGAC 180

554 AAACGGAAGTGGCCAGAGCTTGCAGAGTTTGGAGAAACAGAGATCTCTCATAGTGA 613

181 AAACGGAAGTGGCCAGAGCTTGCAGAGTTTGGAGAAACAGAGATCTCTCATAGTGA 240

614 TTCTCCATCGTCACATACGGAAGAGTGTGTCCTCCCTTGGAGGTCTCTGAAGTTC 673

241 TTCTCCATCGTCACATACGGAAGAGTGTGTCCTCCCTTGGAGGTCTCTGAAGTTC 300

674 TCTAACCTTGGATCTCTTGGCTTCACTGCTGACAGCAGCTTGTAGATGAGCGCTTGAT 733

301 TCTAACCTTGGATCTCTTGGCTTCACTGCTGACAGCAGCTTGTAGATGAGCGCTTGAT 360

734 AAGAGGTGGATGACATGCTTGTGCTGGGCTTGGAGGAAGTATAGAGATTTCACAGA 793

361 AAGAGGTGGATGACATGCTTGTGCTGGGCTTGGAGGAAGTATAGAGATTTCACAGA 420

794 CGCTATATCAGAGATGTTTGGAAATAGCAGGACATCATACATGATATCTTCCAA 853

421 CGCTATATCAGAGATGTTTGGAAATAGCAGGACATCATACATGATATCTTCCAA 480

854 TCAATTGGCTCAAGGAATTTCCAGAGTACCTATCATCTGAGGGAATATGCAACTGGAG 913

481 TCAATTGGCTCAAGGAATTTCCAGAGTACCTATCATCTGAGGGAATATGCAACTGGAG 540

914 ACTAGTACCAAGCTTCTAAAGAAAG-----ACTGGTCCCATTTGTC 939

541 ACTAGTACCAAGCTTCTAAAGAAAGTATTGAGGCTCTGAACACAGTAACTAGACATAT 600

940 -----ACTGGTCCCATTTGTC 955

601 GCCCGGAACAAACCGATGGTTAAACCCGTTTTTTGACAGACCTGTCATTTGTC 660

956 CCCCCTGTCTAGCGCTTGAAGATATGATGTCTCGAAGTGGAGAGATGCTTTTGAA 1015

661 CCCCCTGTCTAGCGCTTGAAGATATGATGTCTCGAAGTGGAGAGATGCTTTTGAA 720

1016 CCGGCTCTGAATCGTGGCAAGTTTCATCCAGGGCCACAGAGCTCAGCCATCCAAAT 1075

721 CCGGCTCTGAATCGTGGCAAGTTTCATCCAGGGCCACAGAGCTCAGCCATCCAAAT 780

1076 AAGATGCCATATACATGAGCTGAGAACAGAGAAGTTATCCCTGTGACCTGTGTGAT 1135

781 AAGATGCCATATACATGAGCTGAGAACAGAGAAGTTATCCCTGTGACCTGTGTGAT 840

1136 CGAATCATCATTTGGGATCGCGAATGGGACGCCACATTAATCCAAATCCACTTGAAC 1195

841 CGAATCATCATTTGGGATCGCGAATGGGACGCCACATTAATCCAAATCCACTTGAAC 900

1196 CAATGAGAAAGAAAGAGATGGATGGATCGATGCTGTCACACCATATGAAATTCAGAT 1255

901 CAATGAGAAAGAAAGAGATGGATGGATCGATGCTGTCACACCATATGAAATTCAGAT 960

1256 GTTTCGCCAGACTATACAAAGAACTTAAGGGAAGGATCCCGAGGGCGAATGATCAA 1315

961 GTTTCGCCAGACTATACAAAGAACTTAAGGGAAGGATCCCGAGGGCGAATGATCAA 1020

1316 GACCTGAATATGACGCTTTAAGACATGTCACGTGGCTTTGGAAGTGTGGGATC 1375

1021 GACCTGAATATGACGCTTTAAGACATGTCACGTGGCTTTGGAAGTGTGGGATC 1080

1376 CAGTTGAGGAGGAGGGGTATGTTGTCGCCAGTCCGCAAGAGAGTCTATGGCGAA 1435

1081 CAGTTGAGGAGGAGGGGTATGTTGTCGCCAGTCCGCAAGAGAGTCTATGGCGAA 1140

1436 TTCTCTGCATAGCAGAAAGCTCCACCATTTCTTTGATGATGTTTAAAGTCTCAG 1495

1141 TTCTCTGCATAGCAGAAAGCTCCACCATTTCTTTGATGATGTTTAAAGTCTCAG 1200

1496 TTCTCTGCATAGCAGAAAGCTCCACCATTTCTTTGATGATGTTTAAAGTCTCAG 1555

1201 TTCTCTGCATAGCAGAAAGCTCCACCATTTCTTTGATGATGTTTAAAGTCTCAG 1260

1556 TGAATGATGATGAGAAAGCAATTTTCTTTGTAACCTTAAAGGTTCTATTATTTAAA 1615

1261 TGAATGATGATGAGAAAGCAATTTTCTTTGTAACCTTAAAGGTTCTATTATTTAAA 1320

1616 GCAGCAGAGATTCACATTTTATACATGAGATCTTCTTTGTGTGTAATACAGAGATTG 1675

1321 GCAGCAGAGATTCACATTTTATACATGAGATCTTCTTTGTGTGTAATACAGAGATTG 1380

1676 ACTGCATCCCTTTAAAAAGATTATGTCCCTGACCTGTGGCTTAAATATCTAATTTCC 1735

1381 ACTGCATCCCTTTAAAAAGATTATGTCCCTGACCTGTGGCTTAAATATCTAATTTCC 1439

1736 AGATGCTTTTGTATGATGATGAGATTTTGTGAGCCCATATTTGGAGTCTGATTTGA 1795

1440 AGATGCTTTTGTATGATGATGAGATTTTGTGAGCCCATATTTGGAGTCTGATTTGA 1499

1796 GTAATGCGAGAGAAAGGCCATCTCCATTGATGATTAAGTAAACCAACTAGTTCTCG 1855

1500 GTAATGCGAGAGAAAGGCCATCTCCATTGATGATTAAGTAAACCAACTAGTTCTCA 1559

1856 GAAATTCTACAGAGAGAGGGAATCAGACTGAGAGAGCTGTGACATATGACTTGAAGACC 1915

```

"note"="unnamed protein product"
/codon_start=1
/protein_id="BAA90923.1"
/db_xref="GI:7019915"
/translacion="MGTEKVIDRKEVLEKEDGLVLHRLKSVDPEMAKCLHPHDKRRKY
ARSLQVPEERGETSLEFRLHROTHEEGGAGLGGTAFNSPCILMTHADAVLDELDRK
VDMLLAGLGLLEELRDLDFHRRYRNOKVSNSDYOGSTISIGFKEFHYLITKEGCTLE
TSNOLLKGLGITALQVYTRRYARKONRRYKKNFSLRPGIYVPYVGLKSYDSVKSEESY
LEPILQVTSFTQGHKPTATPIKMPYVDEAKKRSYTHLDCDLRIILIDGREAAHIKSKN
SHLMQLKRRRLSDAVNTIESQSVSDYNAEKKRSPKPKKSGPQGNDELCSV"
BASE COUNT      541 a      337 c      408 g      463 t
ORIGIN

```

Query Match	76.3%	Score 1556.8	DB 9	Length 1749
Best Local Similarity	95.1%	Pred. No. 0		
Matches 1600	Conservative	0	Mismatches 7	Indels 79
				Gaps 2

QY	374	ACAAATTTATCATGTGAATCTCGTCTGGAAAGTTCTTGCAATACCAAGCCCAAGAG	433
Db	1	ACCAATTTATCATGTGAATCTCGTCTGGAAAGTTCTTGCAATACCAAGCCCAAGAG	60
QY	434	ATGGCACTGAGAAAGTATGACCGAAAGTGAGACTTGAAAGAGAGATGCTTTGA	493
Db	61	ATGGCACTGAGAAAGTATGACCGAAAGTGAGAGCTTGAAAGAGAGATGCTTTGA	120
QY	494	CTTCACAAAGCCTTAAGCAGGTGGACCCAGAAATGCTGCCAAGCTGATCACAATGAC	553
Db	121	CTTCACAAAGCCTTAAGCAGGTGGACCCAGAAATGCTGCCAAGCTGATCACAATGAC	180
QY	554	AAACGCAAGTGGCCAGGAGGCTGGACGTTTTTTGAAGAAACGAAATCCTCATGTGTA	613
Db	181	AAACGCAAGTGGCCAGGAGGCTGGACGTTTTTTGAAGAAACGAAATCCTCATGTGTA	240
QY	614	TTTCTCATCGTCAACATACGGAAGAAGTGGTGGCCCTTGGAAGTCTCTGAAGTTC	673
Db	241	TTTCTCATCGTCAACATACGGAAGAAGTGGTGGCCCTTGGAAGTCTCTGAAGTTC	300
QY	674	TCTAACCTTGATCCTTTGGCTTCATCTGTACCGAGCAATTCATGAGGCTTGGAT	733
Db	301	TCTAACCTTGATCCTTTGGCTTCATCTGTACCGAGCAATTCATGAGGCTTGGAT	360
QY	734	AAGAGGTGATGATGATGATGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTCACAG	793
Db	361	AAGAGGTGATGATGATGATGCTGCTGGGCTCTTGAGAGAACTAAGAGATTTTCACAG	420
QY	794	CGCTTAATTCAGAGAAATGTTTCGAAAAATACCGACGACTATCAACATGATTTCTTCAA	853
Db	421	CGCTTAATTCAGAGAAATGTTTCGAAAAATACCGACGACTATCAACATGATTTCTTCAA	480
QY	854	TCAATTTGGCTTAAGGAATTCACAGATACCGATCAGTGGGGAATAGCACTGGAG	913
Db	481	TCAATTTGGCTTAAGGAATTCACAGATACCGATCAGTGGGGAATAGCACTGGAG	540
QY	914	ACTAGTAACCAAGCTTCTTAAGAAAG-----	939
Db	541	ACTAGTAACCAAGCTTCTTAAGAAAGTATTGAGGCTGTGAACAAGTAACTAAGAGATAT	600
QY	940	-----ACCTGGTCCATGTC	955
Db	601	GCCCGGAACAAACGATGGGTTAAAAACCGTTTTTTGAGACAGACCTGGTCCCTTTGTC	660
QY	956	CCCCCTGTCTAATGGTTAGAGGTATCTGATGCTGGAAGTGGGAGAGACTGCTTTGAA	1015
Db	661	CCCCCTGTCTAATGGTTAGAGGTATCTGATGCTGGAAGTGGGAGAGCTGCTTTGAA	720
QY	1016	CTGTGCTTTGAATTCGTCAAAATTTTCATCCAGGGCCCAAGCCCTACAGCACTCTCAATA	1075
Db	721	CTGTGCTTTGAATTCGTCAAAATTTTCATCCAGGGCCCAAGCCCTACAGCACTCTCAATA	780
QY	1076	AAGATGCCATACATGAAGCTGAGAACCAAGAGACTTATCACCTGTGTGACCTCTGTAT	1135
Db	781	AAGATGCCATACATGAAGCTGAGAACCAAGAGACTTATCACCTGTGTGACCTCTGTAT	840

OY	1136	CGAATCAATCATTTGGGGATTCGGGAATGGCAGCGCACATATAAATCCAAATCCCACTTGAAAC	1195
Db	841	CGAAATCAATCATTTGGGGATTCGGGAATGGGCGAGCGCACATAAATCCAAATCCCACTTGAAAC	900
OY	1196	CAACGAAGAAAAAGAAAGAAAGTGGACGTCAAGTGCATGCAACACATATGAAGACGTAGAGT	1255
Db	901	CAACTGAAGAAAGAAAGAAAGATTTGGACTCAGATGCTGTCAACACCATATGAAGAGTGAAGT	960
OY	1256	GTTTCCCGCAGACTATTAACAAGAAAGAACCTTAAAGGGAAGGGATCCCGAGGCGAGAAATGATCA	1315
Db	961	GTTTCCCGCAGACCAATTAACAAGAAAGAACCTTAAAGGAAGGGATCCCGAGGCGAGAAATGATCA	1020
OY	1316	GAGCTGAATATGACAGCGCTTTAAGAGACATGTCCAGTGCCTTTTGAAGAGTGTGGGATC	1375
Db	1021	GAGCTGAATATGACAGCGCTTTAAGAGACATGTCCAGTGCCTTTTGAAGAGTGTGGGATC	1080
OY	1376	CAGTTCAGAGAGGAGGAGGTATGTTTGTCTCCAGTCTGGGCGAAAGAGAGTGCATATCGGAA	1435
Db	1081	CAGTTCAGAGAGGAGGAGGTATGTTTGTCTCCAGTCTGGGCGAAAGAGAGTGCATATCGGAA	1140
OY	1436	TTCTCTGCATAGCAGAAAAAGCTCCACCAATTTCTTTTGAATGTGTTTAAAGTCTCAGC	1495
Db	1141	TTCTCTGCATAGCAGAAAAAGCTCCACCAATTTCTTTTGAATGTGTTTAAAGTCTCAGC	1200
OY	1496	TTCTCTATAATATGAAGAACAGCAGGTCTGTGCAGCTCCTTGTGTGGCTGTATGTCTCGGAAA	1555
Db	1201	TTCTCTATAATATGAAGAACAGCAGGTCTGTGCAGCTCCTTGTGTGGCTGTATGTCTCGGAAA	1260
OY	1556	TGATGTAGTTCAGGAAGAACATTTTCTTTTCTTGAACCTTAAAGCTTATATTTAA	1615
Db	1261	TGATGTATTTCAAGAAAGCATTTTTTTTTTCTTTTAACCTTAAAGGTCTATATTTAA	1320
OY	1616	GCAGCAGAGATTCACATTTTTTATATCATAGAGATCTTCTTGTGTGTAATACAGAGATTG	1675
Db	1321	GCAGCAGCAGATTCACATTTTTTATATCATAGAGATCTTCTTGTGTGTAATACAGAGATTG	1380
OY	1676	ACTGCATCCCTTTAAAAAAGTTTATATGTCCCTGCAGCTCTGGTAAATATATATTTCC	1735
Db	1381	ACTGCATCCCTTTTAAAAAGTTTATATGTCCCTGCAGCTCTGGTAAATATATATTTCC	1439
OY	1736	AGATCTCTTTTGAATGACTGAGTAAGATTTGTGAGCCACATATTTGGAGTTCAGATTTGA	1795
Db	1440	AGATCTCTTTTGAATGACTGAGTAAGATTTGTGAGCCACATATTTGGAGTTCAGATTTGA	1499
OY	1796	GTGAATGCGAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAAACTAGTTCG	1855
Db	1500	GTGAATGCGAGAAAGGCGCATCTCCATTTGAGATGATTAAGTGAACCAAACTAGTTCG	1559
OY	1856	GAAATCTACAGAGAAAGGAGGAATCAGACTAGAGAAAGCTGTGACATTAAGAGCTTGAAGACC	1915
Db	1560	GAAATCTACAGAGAAAGGAGGAATCAGACTAGAGAAAGCTGTGACATTAAGAGCTTGAAGACC	1619
OY	1916	AAAGACTTTGAATTTGGAGCTGCTCATGTGTGAGTTATATCACTCTGCTCTTCTAT	1975
Db	1620	AAAGACTTTGAATTTGGAGCTGCTCATGTGTGAGTTATATCACTCTGCTCTTCTAT	1679
OY	1976	TGAGTTACAAATCTATATTTTATTTGAAGTTTAAATTAAGAAAAAATTTACACAAGAAAA	2035
Db	1680	TGAGTTACAAATCTATATTTTATTTGAAGTTTAAATTAAGAAAAAATTTAAAAAATAA	1739
OY	2036	AAAAAA 2041	
Db	1740	AAAAAA 1745	

RESULT_5	LOCUS	DEFINITION	ACCESSION	VERSION
AK074222	1749 bp	mRNA	AK074222	GI:18676765
		linear		
		PRI 15-FEB-2002		
		Homo sapiens CDNA FLJ23642 fis, clone COL01371, highly similar to		
		Homo sapiens tRNA isopentenylpyrophosphate transferase precursor		
		tRNA.		

QY 661 TCCCTGAAAGTCTCTAACCCCTGCATCCCTGGCTTCAGCTGACAGGACAGTCTAGA 720  
 DB 665 TCCCTGAAAGTCTCTAACCCCTGCATCCCTGGCTTCAGCTGACAGGACAGTCTAGA 724  
 QY 721 TGAGCCCTGGATAGAGGGTGGATGACATGCTGCTGGCTGGCTGGAGAGACTAG 780  
 DB 725 TGAGCCCTGGATAGAGGGTGGATGACATGCTGCTGGCTGGCTGGAGAGACTAG 784  
 QY 781 AGATTTTTCACAGACGCTATATACAGAAAGTGTTCGGAAAAAT- 823  
 DB 785 AGATTTTTCACAGACGCTATATACAGAAAGTGTTCGGAAAAATAGATGAATGGCTCAAC 844  
 QY 824 ----- 823  
 DB 845 TAGGTCACATCCAAAGCCCTGCTAGCTTTAGAGACAAATGGCTTATCTGTGATACC 904  
 QY 824 ----- AGCCAGAGACTATACATGATGATCTTCACATCAATGGCTTCAAGAAAT 873  
 DB 905 TCCCTGTTCCAGCAGGACGATATACATGATGATCTTCACATCAATGGCTTCAAGAAAT 964  
 QY 874 TCACAGATCTGATCTACTGAGGAAAAATGACACTGGAGACTATGAACAGCTTCTAAA 933  
 DB 965 TCACAGATCTGATCTACTGAGGAAAAATGACACTGGAGACTATGAACAGCTTCTAAA 1024  
 QY 934 GAAAG----- 938  
 DB 1025 GAAAGGATTTGAGGCTCTGAAACAGTAACTAGAGATATGCCGGAACAAACCGATG 1084  
 QY 939 ----- GACCTGGTCCCATTTGCCCTGCTCCCTGCTATGAGCTAGA 975  
 DB 1085 GGTAAAAACCGTTTTTTTGTAGCAGACCTGGTCCCATTTGCCCTGCTATGAGCTAGA 1144  
 QY 976 GGTATCTGATGCTCTGAAAGTGGAGAGAGTGTCTTCTGAACTGCTCTTGAATCTGCA 1035  
 DB 1145 GGTATCTGATGCTCTGAAAGTGGAGAGAGTGTCTTCTGAACTGCTCTTGAATCTGCA 1204  
 QY 1036 AAGTTTCATCCAGGCGCCACAAAGCTTACAGGCTCTCCATTAAGATGCCATPACATGAGC 1095  
 DB 1205 AAGTTTCATCCAGGCGCCACAAAGCTTACAGGCTCTCCATTAAGATGCCATPACATGAGC 1264  
 QY 1096 TGAACAACAAGAAATATACCTGCTGAGCTGCTGTGATGCAATCATCTGGGGGATCG 1155  
 DB 1265 TGAACAACAAGAAATATACCTGCTGAGCTGCTGTGATGCAATCATCTGGGGGATCG 1324  
 QY 1156 CGAATGGGACGCGACATATAATCCAAATCCACTTGAACCAACTGAAGAAAAAGAGAG 1215  
 DB 1325 CGAATGGGACGCGACATATAATCCAAATCCACTTGAACCAACTGAAGAAAAAGAGAG 1384  
 QY 1216 ATTGGACTCAGATGCTGTCTCAACACCATAGAAAGTGTTCCTCCAGACTATACAA 1275  
 DB 1385 ATTGGACTCAGATGCTGTCTCAACACCATAGAAAGTGTTCCTCCAGACTATACAA 1444  
 QY 1276 AGAACCTAAAGGAGAGGATCCCGAGGCGAATGATCAAGAGTGAAGAAAGCGCTTTA 1335  
 DB 1445 AGAACCTAAAGGAGAGGATCCCGAGGCGAATGATCAAGAGTGAAGAAAGCGCTTTA 1504  
 QY 1336 AGAGACATGTCAGATGCGCTTTGAAAAGTGTGGGGGATCCAGTTCAGAGAGGAGGTA 1395  
 DB 1505 AGAGACATGTCAGATGCGCTTTGAAAAGTGTGGGGGATCCAGTTCAGAGAGGAGGTA 1564  
 QY 1396 TGTGTTGCTCCCACTGCGGCAAAAGAGTGTCTATGCGGAATTTCTCTCATAGCAGAAAAG 1455  
 DB 1565 TGTGTTGCTCCCACTGCGGCAAAAGAGTGTCTATGCGGAATTTCTCTCATAGCAGAAAAG 1624  
 QY 1456 CTCACACATTTTCTTTGATGTTGTTTAAAGTCTACGTTCTCTATATAGAAACAGC 1515  
 DB 1625 CTCACACATTTTCTTTGATGTTGTTTAAAGTCTACGTTCTCTATATAGAAACAGC 1684  
 QY 1516 AGGCTTTGTGAGCTCTTGTGTGCTGATGTCTCTGAAAATGATGATGTTACAGAAAAGCA 1575  
 DB 1685 AGGCTTTGTGAGCTCTTGTGTGCTGATGTCTCTGAAAATGATGATGTTACAGAAAAGCA 1744  
 QY 1576 TTTTCTTTTCTTTGAAACCTTAAAGTCTTATATATAAAGCAGACAGATTCACACATTT 1635

DB 1745 TTTTCTTTTCTTTGAAACCTTAAAGTCTTATATATAAAGCAGACAGATTCACATTT 1804  
 QY 1636 TTATACATGAGATCTCTTGTGTGTAATACAGAGATGATGATGATCCCTTTAAAGAA 1695  
 DB 1805 TTATACATGAGATCTCTTGTGTGTAATACAGAGATGATGATGATCCCTTTAAAGAA 1863  
 QY 1696 GTTTTATGTCCTGACTCTGGCTAAATATCTAATTTTCCAGATGCTTTTGTGATGACT 1755  
 DB 1864 GTTTTATGTCCTGACTCTGGCTAAATATCTAATTTTCCAGATGCTTTTGTGATGACT 1923  
 QY 1756 GAAGTATTTGAGCCACATATTGGGAGTCTGATTTGATGATGATGAGCAGAAAGGCC 1815  
 DB 1924 GAAGTATTTGAGCCACATATTGGGAGTCTGATTTGATGATGATGAGCAGAAAGGCC 1983  
 QY 1816 ATCTCATTTGATGATTTAAGTAAAGCAACTAGTTCTCGAAATTTACAGAGAGAGG 1875  
 DB 1984 ATCTCATTTGATGATTTAAGTAAAGCAACTAGTTCTCGAAATTTACAGAGAGAGG 2043  
 QY 1876 GAATCAGACTGAGAAAGCTGTGACATGACACTTGAAGACCAAGACTTTGAAATTTGCCA 1935  
 DB 2044 GAATCAGACTGAGAAAGCTGTGACATGACACTTGAAGACCAAGACTTTGAAATTTGCCA 2103  
 QY 1936 GCTGCTCATGCTGAGTATATACACGCTGCTCTTCTATGATGATTAATGATTT 1995  
 DB 2104 GCTGCTCATGCTGAGTATATACACGCTGCTCTTCTATGATGATTAATGATTT 2163  
 QY 1996 TTATTAAGTTTAAATTAAGAAAAATTTACAGAAAAAATAAATAA 2041  
 DB 2164 TTATTAAGTTTAAATTAAGAAAAATTTAAATTAATTAATTAATTAATTAATTT 2209

RESULT 4  
 AK000068 1749 bp mRNA linear PRI 22-FEB-2000  
 LOCUS  
 DEFINITION Homo sapiens cDNA FLJ20061 fis, clone COL01383.  
 ACCESSION AK000068  
 VERSION AK000068.1 GI:7019914  
 KEYWORDS oligo capping; fis (full insert sequence);  
 SOURCE Homo sapiens colon cDNA clone COL01383.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE  
 1 (sites)  
 Kawabata, A., Hiki, J., Kobatake, N., Inagaki, H., Ikema, Y.,  
 Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,  
 Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
 NEDO human cDNA sequencing project  
 UNPUBLISHED  
 2 (bases 1 to 1749)  
 Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,  
 Shibahara, T., Tanaka, T. and Nakamura, Y.  
 Direct Submision  
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,  
 University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,  
 Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna@ims.u-tokyo.ac.jp,  
 Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
 COMMENT  
 NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan: cDNA full insert  
 sequencing; Research Association for Biotechnology: cDNA library  
 construction; 5'-3'-end one pass sequencing; Department of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).  
 FEATURES  
 1. 1749  
 Location/Qualifiers  
 source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="COL01383"  
 /tissue\_type="colon"  
 /clone\_lib="COL"  
 /note="cloning vector pME18SFL3"  
 CDS  
 61..1041

QY	1430	GCAGAAATTCCTCGATATGACGAAAGAAAGCTCCACACATTTTCTTTGATGATGTTTTAAAGT	1449
Db	1520	GCAGAAATTCCTCGATATGACGAAAGAAAGCTCCACACATTTTCTTTGATGATGTTTTAAAGT	1579
QY	1490	CTCACGTTCTCTATAAATAAGAAACAGCAGCTTGTCTCAGCTCCTGTGTGCGTGTGATGTGC	1549
Db	1580	CTCACGTTCTCTATAAATAAGAAACAGCAGCTTGTCTCAGCTCCTGTGTGCGTGTGATGTGC	1639
QY	1550	TGGAATGATGCTAGTCTCAGGAAAGCAATTTTTTTTTTTCTTTGAACTTAAAGTTCTATA	1609
Db	1640	TGGAATGATGCTAGTCTCAGGAAAGCAATTTTTTTTTTTCTTTGAACTTAAAGTTCTATA	1699
QY	1610	TTAAAGACAGACAGATTTCCACATTTTATATCATGAGAGATCTCTTGTGTGTAATACA	1669
Db	1700	TTAAAGACAGACAGATTTCCACATTTTATATCATGAGAGATCTCTTGTGTGTAATACA	1759
QY	1670	GGATTGACTGCATCCCTTTAAAGAAAGTTTATGTCCCTGACTGTGGCTAAATATATCTA	1729
Db	1760	GGATTGACTGCATCCCTTT-AAAGAAAGTTTATGTCCCTGACTGTGGCTAAATATATCTA	1818
QY	1730	ATTTCACAGATCTTTTGTGAATGACTGAAAGTATTTGTGACCAATATTTGGAGTTCTAG	1789
Db	1819	ATTTCACAGATCTTTTGTGAATGACTGAAAGTATTTGTGACCAATATTTGGAGTTCTAG	1878
QY	1790	ATTGAGAGTATGACAGAAAGGCGCANTCTCATTTAGATGATTAAAGAACCAAACTAG	1849
Db	1879	ATTGAGAGTATGACAGAAAGGCGCANTCTCATTTAGATGATTAAAGAACCAAACTAG	1938
QY	1850	TTCTCGGAATTCACAGAGAAAGAGGAATCAGACTGAGAGAGCTGTACATATAGACTTG	1909
Db	1939	TTCTCGGAATTCACAGAGAAAGAGGAATCAGACTGAGAGAGCTGTACATATAGACTTG	1998
QY	1910	AAGACCAAAAGCTTGAATTTTGGAGCTGCTCATGTGTGATTTATATCATCTGCTCT	1969
Db	1999	AAGACCAAAAGCTTGAATTTTGGAGCTGCTCATGTGTGATTTATATCATCTGCTCT	2058
QY	1970	TTCTATGAGTTTAAATATCTATATTTTATTTGAAGTTTAAATTAAGAAAAATTTACAG	2029
Db	2059	TTCTATGAGTTTAAATATCTATATTTTATTTGAAGTTTAAATTAAGAAAAATTTACAG	2118
QY	2030	AAAAAAAAAAAAA 2041	
Db	2119	AAAAAAAAAAAAA 2130	
RESULT 3			
LOCUS	AF074918	2216 bp	RNA
DEFINITION	Homo sapiens tRNA Isopentenylpyrophosphate transferase precursor		
ACCESSION	AF074918		
VERSION	AF074918.1		
KEYWORDS	GI:11137964		
SOURCE			
ORGANISM	Homo sapiens.		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE			
AUTHORS	Golovko, A., Hjalms, G., Stihon, F. and Nicander, B.		
TITLE	Cloning of a human tRNA Isopentenyl transferase		
JOURNAL	Gene 258 (1-2), 85-93 (2000)		
MEDLINE	20564178		
PUBMED	11111046		
REFERENCE	2 (bases 1 to 2216)		
AUTHORS	Golovko, A. and Hjalms, G.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-JUN-1998) Dep. of Plant Biology, Swedish University		
	of Agricultural Sciences, Box 7080, SE-750 07 uppsala, Sweden		
FEATURES			
source	1. 2216		
	/organism="Homo sapiens"		
	/db_xref="dbEST:AA332152"		

	db_xref="taxon:9606"	/note="extension of sequence by 5' RACE"
precursor_RNA	1..2216	/note="incompletely processed mRNA"
CDS	join(15..829,917..1505)	/note="lipoase"
	/codon_start=1	/product="tRNA Isopentenyl pyrophosphate transferase"
	/protein_id="AA031324.1"	/db_xref="GI:11137965"
intron	830..916	/evidence="experimental"
polyA_signal	2177..2182	
BASE COUNT	651 a 461 c 534 g 570 t	
ORIGIN		
Query Match	89.9%; Score 1834.4; DB 9; Length 2216;	
Best Local Similarity	92.2%; Pred. No. 0;	
Matches 2034; Conservative	0; Mismatches 6; Indels 166; Gaps 3;	
1	CTGCATAGATGCGCTCGTGCGGCTGCACGACAGTCTCTGTGGCAGTGGCTCAG	60
5	CTGCATAGATGCGCTCGTGCGGCTGCACGACAGTCTCTGTGGCAGTGGCTCAG	64
61	GGGCTGCAACGACCGCTCACTCTGTGTGTGTCTCGGGGGCAGGGGACCGGCAATC	120
65	GGGCTGCAACGACCGCTCACTCTGTGTGTGTCTCGGGGGCAGGGGACCGGCAATC	124
121	CACGCTGGCGTGTGACGTAGGCGCAGCGCTCGCGGCTGAGATCGTACGCTGCTCAT	180
125	CACGCTGGCGTGTGACGTAGGCGCAGCGCTCGCGGCTGAGATCGTACGCTGCTCAT	184
181	GCAGGCTGTATGAGGCTGTACATATCACCACAAAGTTTCTGCCAAGACAGAGAT	240
185	GCAGGCTGTATGAGGCTGTACATATCACCACAAAGTTTCTGCCAAGACAGAGAT	244
241	CTGCCGGGACCCAGATAGTACGCTTGTGGATCCTCTGTGACCAATTAACAGTGTGA	300
245	CTGCCGGGACCCAGATAGTACGCTTGTGGATCCTCTGTGACCAATTAACAGTGTGA	304
301	CTTCGAAATATAGGACACTGCTGTATGTAAGATATATTGGCCGAGACAAATTCCTAT	360
305	CTTCGAAATATAGGACACTGCTGTATGTAAGATATATTGGCCGAGACAAATTCCTAT	364
361	TGTTGTGGAGAGAACCAATATTACATTTGAATCTCTGCTGTGAAAGTCTGTCAATAC	420
365	TGTTGTGGAGAGAACCAATATTACATTTGAATCTCTGCTGTGAAAGTCTGTCAATAC	424
421	CAAGCCCCAGAGATGGGCACTGAGAAATGATTGACCCGAAAGTGGAGCTTGAAAGGA	480
425	CAAGCCCCAGAGATGGGCACTGAGAAATGATTGACCCGAAAGTGGAGCTTGAAAGGA	484
481	GGATGCTCTTACTTACAAAGCGCTTAAGCCAGGTGAGCCAGAAATGCTCCAGCT	540
485	GGATGCTCTTACTTACAAAGCGCTTAAGCCAGGTGAGCCAGAAATGCTCCAGCT	544
541	GCATCCACATGACAAACGAAAGTGGCCGAGGCTTGAAGTTTGAAGAAACAGGAT	600
545	GCATCCACATGACAAACGAAAGTGGCCGAGGCTTGAAGTTTGAAGAAACAGGAT	604
601	CTCTATAGTGAATTTCTCATGTCACATACGAAAGAGAGTGTGTCCTCTGGAG	660
605	CTCTATAGTGAATTTCTCATGTCACATACGAAAGAGAGTGTGTCCTCTGGAG	664

```

QY      2041 A 2041
Db      2041 A 2041

RESULT 2
LOCUS   AX405704      2130 bp      DNA      linear      PAT 14-JUN-2002
DEFINITION Sequence 119 from Patent WO0222660.
ACCESSION AX405704
VERSION  AX405704.1 GI:21438847
KEYWORDS
SOURCE   human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
  1. Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
    Xue, A.J., Yang, Y., Wehrman, F. and Dimanac, R.T.
    Novel nucleic acids and polypeptides
    Patent: WO 0222660-A 119 21-MAR-2002;
    HYSEQ, INC. (US)

FEATURES
    source             1..2130
                        Location/Qualifiers
                        1..2130
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        191..1426
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAD34811.1"
                        /db_xref="GI:21438848"
                        /translation="MOVVEGLDIITNKVSAOERICRHHMISFVDELYNTVYDFNR
                        RAFLAIEDIFARDKIPVVGNTNYIESILMKVLVTKFOENGTEKVIDRKVELEKED
                        GLVLRLSQVDPPEMAKLPHPDKRVASLQVFEETGSHSEFLRHOTEEGGPBG
                        GPLKSNPCILWLDADQVLEDRDLKRVMDLAAGLEELKQVTRAKRKNRYKNR
                        YOHGIFOSIGKEFEHEHYLTTECKTLETSNOLKLSIELKQVTRAKRKNRYKNR
                        ELSRPPRIYPYVGLSEVSDVSWBESVLEPALEIYOSFIOGHRPTATPIKMPINEN
                        KRSHLDCDLDRIIIGDREMAHIIKSKHNLKRRRLDSAVNTIESVSPHNK
                        EPKEKSGPQNOELKCSY"

BASE COUNT      620 a      439 c      522 g      549 t
ORIGIN
Query Match      94.1%; Score 1921.2; DB 6; Length 2130;
Best Local Similarity 95.9%; Pred. No. 0;
Matches 2025; Conservative 0; Mismatches 8; Indels 79; Gaps 2;

```

```

Db      380 GGAAGAACCAATTATATGATTCCTCTGGAAGTTCTTCTCAATACCAAGCC 439
QY      428 CAGAGATGGGCACTGAGAAAGTATGACGAAAGTGGAGCTTGAAGAGAGATG 487
Db      440 CAGAGATGGGCACTGAGAAAGTATGACGAAAGTGGAGCTTGAAGAGAGATG 499
QY      488 CTGTACTTACAAAGCCCTTAAGCCAGGTGGACCCAGAAATGGCTGCCAAGCTG 547
Db      500 CTGTACTTACAAAGCCCTTAAGCCAGGTGGACCCAGAAATGGCTGCCAAGCTG 559
QY      548 CATGACAAACGCAAGAGGGCCAGAGCTTCAAGTTTGAAGAAACAGAAATCTCAT 607
Db      560 CATGACAAACGCAAGAGGGCCAGAGCTTCAAGTTTGAAGAAACAGAAATCTCAT 619
QY      608 AGTGAATTTCTCATGCTCAACATACGAAAGAGTGGTCCCTTGGAGGCTCTCTG 667
Db      620 AGTGAATTTCTCATGCTCAACATACGAAAGAGTGGTCCCTTGGAGGCTCTCTG 679
QY      668 AAGTCTCTAACCCCTTGATCTCTTGGCTTCATGCTGACCAAGCAGTTTATGAGCC 727
Db      680 AAGTCTCTAACCCCTTGATCTCTTGGCTTCATGCTGACCAAGCAGTTTATGAGCC 739
QY      728 TTGGATTAAGAGGGGATGATGATGCTGCTGGGCTTTGGAGAACTAAGATTTT 787
Db      740 TTGGATTAAGAGGGGATGATGATGCTGCTGGGCTTTGGAGAACTAAGATTTT 799
QY      788 CACAGAGCTATATATCAGAAAGTGTTCGAAATATGCCAGACTATCAACATGATATC 847
Db      800 CACAGAGCTATATATCAGAAAGTGTTCGAAATATGCCAGACTATCAACATGATATC 859
QY      848 TTCCATCAATTTGGCTTCAAGAAATTTACAGAGTACTGATCACTGAGGAAATGCA 907
Db      860 TTCCATCAATTTGGCTTCAAGAAATTTACAGAGTACTGATCACTGAGGAAATGCA 919
QY      908 CTGAGAGCTATGTAACAGCTTCTTAAGAAG----- 938
Db      920 CTGAGAGCTATGTAACAGCTTCTTAAGAAGATTTAGAGCTCTGAACAAGTACTAAG 979
QY      939 -----GACCTGGTCCC 949
Db      980 AGATATGCCCGGAACAAACCGATGGGTTAAAAACCGTTTTTTGACGACAGCTGGTCCC 1039
QY      950 ATTGTCCCCCTGTCTATGCTTATGAGATATGATGATCTTCGAAGTGGAGAGTCTGTT 1009
Db      1040 ATTGTCCCCCTGTCTATGCTTATGAGATATGATGATCTTCGAAGTGGAGAGTCTGTT 1099
QY      1010 CTGGAACCTGCTTGAAGTGTGCAAGTTTATCCAGGGCCCAAGAGCTTACAGCCACT 1069
Db      1100 CTGGAACCTGCTTGAAGTGTGCAAGTTTATCCAGGGCCCAAGAGCTTACAGCCACT 1159
QY      1070 CCAATAAAGATGCAATATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1129
Db      1160 CCAATAAAGATGCAATATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1219
QY      1130 TGTGATGAAATCATCATTTGGGATTCGGAATGGGCAAGCAATTAATCCAAATCCAC 1189
Db      1220 TGTGATGAAATCATCATTTGGGATTCGGAATGGGCAAGCAATTAATCCAAATCCAC 1279
QY      1190 TTGAACCACTGAAGAAAGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1249
Db      1280 TTGAACCACTGAAGAAAGAAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1339
QY      1250 CAGAGTGTTCCTCCAGACTTAAACAAAGAACTTAAGAGGAAAGGATCCCGAGGGCAGAAAT 1309
Db      1340 CAGAGTGTTCCTCCAGACTTAAACAAAGAACTTAAGAGGAAAGGATCCCGAGGGCAGAAAT 1399
QY      1310 GATCAAGAGCTGAATATGACAGCTTTAAGAGACATGTCAGTGGCTTTGGAAAGTGCTG 1369
Db      1400 GATCAAGAGCTGAATATGACAGCTTTAAGAGACATGTCAGTGGCTTTGGAAAGTGCTG 1459
QY      1370 GGGATCACTTCCAGAGGAGGAGGAGGATGTTGCTCCAGTCTGGGCAAAAGAGTCTAT 1429
Db      1460 GGGATCACTTCCAGAGGAGGAGGAGGATGTTGCTCCAGTCTGGGCAAAAGAGTCTAT 1519

```

FEATURES		Location/Qualifiers	
Source		1..2041	
		/organism="unidentified"	
		/db_xref="taxon:32644"	
BASE COUNT	589 a	421 c	502 g 529 t
ORIGIN			
Query Match 100.0%; Score 2041; DB 6; Length 2041;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2041; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CTGCCTAAGATGGCGTCCGTGGCGCTGCACAGACAGTCTCTGGGAGTGGGCTCG	60
DB	1	CTGCCTAAGATGGCGTCCGTGGCGCTGCACAGACAGTCTCTGGGAGTGGGCTCG	60
QY	61	GGCGCTGCAAGGACCCCTACCTCTGTAGTGTCTCGGGGGCAGGGGACGGGCAATC	120
DB	61	GGCGCTGCAAGGACCCCTACCTCTGTAGTGTCTCGGGGGCAGGGGACGGGCAATC	120
QY	121	CACGCTGGCGTTCAGCTAGGCGCAGCGCTCGCGGTGAGATCTCAGCGCTGACTCAT	180
DB	121	CACGCTGGCGTTCAGCTAGGCGCAGCGCTCGCGGTGAGATCTCAGCGCTGACTCAT	180
QY	181	GCAGGTCTATGAAGGCTTAGACATCATCAACAAGTTTCTGCCAAGACAGAGAT	240
DB	181	GCAGGTCTATGAAGGCTTAGACATCATCAACAAGTTTCTGCCAAGACAGAGAT	240
QY	241	CTGGCGGACACATGATAGCTTTGTGATCTCTGTGACCAATTACAGAGTGTGA	300
DB	241	CTGGCGGACACATGATAGCTTTGTGATCTCTGTGACCAATTACAGAGTGTGA	300
QY	301	CTTCAGAAATAGAGCAACTCTCTGATTAAGATATATTTGCCGAGCAAAATTCCTAT	360
DB	301	CTTCAGAAATAGAGCAACTCTCTGATTAAGATATATTTGCCGAGCAAAATTCCTAT	360
QY	361	TTTGTGGGAGAACCAATTATTACATTAATCTCTGCTCTGGAAAGTCTTGTCAATAC	420
DB	361	TTTGTGGGAGAACCAATTATTACATTAATCTCTGCTCTGGAAAGTCTTGTCAATAC	420
QY	421	CAAGCCCCAGAGATGGGCACTGAGAAATGATTGACGAAAGTGGAGCTGAAAGAA	480
DB	421	CAAGCCCCAGAGATGGGCACTGAGAAATGATTGACGAAAGTGGAGCTGAAAGAA	480
QY	481	GGATGGCTTTGACTTACAAAAGCCCTAAGCCAGGTGACCCAGAAATGGCTGCCAAGT	540
DB	481	GGATGGCTTTGACTTACAAAAGCCCTAAGCCAGGTGACCCAGAAATGGCTGCCAAGT	540
QY	541	GCATCCACATGACAAAGGAAAGTGGCGCAGAGCTTGCAGAGTTTGAAGAAACAGGAAT	600
DB	541	GCATCCACATGACAAAGGAAAGTGGCGCAGAGCTTGCAGAGTTTGAAGAAACAGGAAT	600
QY	601	CTCTCATAGTAAATTTCTCATCTGTCACATACAGGAAGAGTGGTGGCTTGGAGG	660
DB	601	CTCTCATAGTAAATTTCTCATCTGTCACATACAGGAAGAGTGGTGGCTTGGAGG	660
QY	661	TCCTCTGAAGTTCTCTAACCCCTTGACCTTTGGCTTCACTGACAGGCAAGTTCTAGA	720
DB	661	TCCTCTGAAGTTCTCTAACCCCTTGACCTTTGGCTTCACTGACAGGCAAGTTCTAGA	720
QY	721	TCAGGCGCTGATTAAGAGGTTGATGATCTGCTGCTGGGCTTGGAGAACTAAG	780
DB	721	TCAGGCGCTGATTAAGAGGTTGATGATCTGCTGCTGGGCTTGGAGAACTAAG	780
QY	781	AGATTTTACAGACGCTATTAATCAGAAAGTGTTCGAAAAATAGCAGAGACTATCAACA	840
DB	781	AGATTTTACAGACGCTATTAATCAGAAAGTGTTCGAAAAATAGCAGAGACTATCAACA	840
QY	841	TGGTATCTTCCAAATCAATTGGGCTCAAGGAATTTCAAGAGTACCTGATCACTGAGGAAA	900
DB	841	TGGTATCTTCCAAATCAATTGGGCTCAAGGAATTTCAAGAGTACCTGATCACTGAGGAAA	900
QY	901	ATGCACACTGAGACTAGTAACCAAGCTCTTAAGAAAGAGCACTGGTCCATTGTCCTCCC	960
DB	901	ATGCACACTGAGACTAGTAACCAAGCTCTTAAGAAAGAGCACTGGTCCATTGTCCTCCC	960
QY	961	TGCTATGAGCTTAGAGTATCTGATGTCTCGAAGTGGAGAGAGTGTCTTGAACCTGC	1020
DB	961	TGCTATGAGCTTAGAGTATCTGATGTCTCGAAGTGGAGAGAGTGTCTTGAACCTGC	1020
QY	1021	TCCTGAATTCGTGCAAGTTTCATCCAGGCGCACAGGCTTACAGCACTCCAAATTAAGAT	1080
DB	1021	TCCTGAATTCGTGCAAGTTTCATCCAGGCGCACAGGCTTACAGCACTCCAAATTAAGAT	1080
QY	1081	GCCATACATGAGCTGAGAACAAAGAAAGTATACCTGTGTGACCTGTGATCGAAT	1140
DB	1081	GCCATACATGAGCTGAGAACAAAGAAAGTATACCTGTGTGACCTGTGATCGAAT	1140
QY	1141	CATCATTTGGGATTCGCGAATGGGCGACGCGACATTAATTCCTCACTTGAACCACT	1200
DB	1141	CATCATTTGGGATTCGCGAATGGGCGACGCGACATTAATTCCTCACTTGAACCACT	1200
QY	1201	GAGAAAGAGAGAGATTGGACTGATGCTGTCAACACCATAGAAAGTCAAGTGTTC	1260
DB	1201	GAGAAAGAGAGAGATTGGACTGATGCTGTCAACACCATAGAAAGTCAAGTGTTC	1260
QY	1261	CCCAGACTTACAAAGAACTTAAGGAGAGGATCCCGAGGCGCAATGATCAAGAGCT	1320
DB	1261	CCCAGACTTACAAAGAACTTAAGGAGAGGATCCCGAGGCGCAATGATCAAGAGCT	1320
QY	1321	GAATGACGCGTTTAAGAGACATGTCCAGTGGCTTTGGAAGAGTGGGAGTCCAGTT	1380
DB	1321	GAATGACGCGTTTAAGAGACATGTCCAGTGGCTTTGGAAGAGTGGGAGTCCAGTT	1380
QY	1381	CAGAGGAGGAGGATATGTTGTCTCCAGTCTGGGCAAGAGTCTATGGGAATCTC	1440
DB	1381	CAGAGGAGGAGGATATGTTGTCTCCAGTCTGGGCAAGAGTCTATGGGAATCTC	1440
QY	1441	TGCTATGACGAAAGAGTCCACATTTCTTTGATGAGTGTAAAGCTCACGTTCTC	1500
DB	1441	TGCTATGACGAAAGAGTCCACATTTCTTTGATGAGTGTAAAGCTCACGTTCTC	1500
QY	1501	TATTAATAGAAACAGAGCTTGTGACAGCTCTGTGTGTGATGATGTCTGAAATGATG	1560
DB	1501	TATTAATAGAAACAGAGCTTGTGACAGCTCTGTGTGTGATGATGTCTGAAATGATG	1560
QY	1561	TAGTTGAGAAAGCATTTTTTTTCTTTGAACTTTAAAGTCTTATTTAAAGCAGC	1620
DB	1561	TAGTTGAGAAAGCATTTTTTTTCTTTGAACTTTAAAGTCTTATTTAAAGCAGC	1620
QY	1621	ACAGATTCACATTTTATACATGAGGATCTTGTGTGTGATGATGATGATGATGATG	1680
DB	1621	ACAGATTCACATTTTATACATGAGGATCTTGTGTGTGATGATGATGATGATGATG	1680
QY	1681	ATCCCTTTAAAGAGTTTATGTCCTGACTCTGCTGCTAAATTTCTAATTTCCAGATG	1740
DB	1681	ATCCCTTTAAAGAGTTTATGTCCTGACTCTGCTGCTAAATTTCTAATTTCCAGATG	1740
QY	1741	CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1800
DB	1741	CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1800
QY	1801	TGGAGGAAGAGGCGCATCCCATTTGAGATGATTAAGTGAACCAACTGATTCGGAAT	1860
DB	1801	TGGAGGAAGAGGCGCATCCCATTTGAGATGATTAAGTGAACCAACTGATTCGGAAT	1860
QY	1861	CTACAGAGAGAGGAGATCAGACTGAGAGAGCTGACATAGGACTTGAAGACCAAGA	1920
DB	1861	CTACAGAGAGAGGAGATCAGACTGAGAGAGCTGACATAGGACTTGAAGACCAAGA	1920
QY	1921	CTTTGAATTTTGGAGCTCTCATGTGTGATTTATATACATGCTGTCTTTTATTTGAGT	1980
DB	1921	CTTTGAATTTTGGAGCTCTCATGTGTGATTTATATACATGCTGTCTTTTATTTGAGT	1980
QY	1981	TACAAATCTAATTTTATTTATTTGAGTTTAAATTAAGAAAAATTTCAAGAAAAAAA	2040
DB	1981	TACAAATCTAATTTTATTTATTTGAGTTTAAATTAAGAAAAATTTCAAGAAAAAAA	2040



Gencore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:07:03 ; Search time 6306.89 Seconds

(without alignments)  
9418.083 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 2041

Sequence: 1 CTGCCNTAAGATGCGCTCG.....TTTACAGAAAAA 2041

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenBank: 1: gb\_ba: 2: gb\_ntg: 3: gb\_in: 4: gb\_om: 5: gb\_ov: 6: gb\_pat: 7: gb\_ph: 8: gb\_pl: 9: gb\_pr: 10: gb\_ro: 11: gb\_sts: 12: gb\_sy: 13: gb\_un: 14: gb\_vl: 15: em\_ba: 16: em\_fun: 17: em\_hum: 18: em\_in: 19: em\_mu: 20: em\_om: 21: em\_ot: 22: em\_ov: 23: em\_pat: 24: em\_ph: 25: em\_pl: 26: em\_ro: 27: em\_sts: 28: em\_un: 29: em\_vl: 30: em\_htg\_hum: 31: em\_htg\_in: 32: em\_htg\_other: 33: em\_htg\_mus: 34: em\_htg\_pin: 35: em\_htg\_rod: 36: em\_htg\_mam: 37: em\_htg\_vrt: 38: em\_sy: 39: em\_hcgo\_hum: 40: em\_hcgo\_mus: 41: em\_hcgo\_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2041	100.0	2041	6 A98668	A98668 Sequence 3
2	1921.2	94.1	2130	6 AX405704	AX405704 Sequence
3	1834.4	89.9	2216	9 AF074918	AF074918 Homo sapi
4	1556.8	76.3	1749	9 AK000068	AK000068 Homo sapi
5	1556.8	76.3	1749	9 AK004222	AK004222 Homo sapi
6	1489	73.0	1844	9 BC010741	BC010741 Homo sapi
7	1306.4	64.0	1308	9 AY052768	AY052768 Homo sapi
8	1119.4	54.8	2045	10 BC019812	BC019812 Homo sapi
9	850.4	41.7	166496	9 HS118721	HS118721 Human DNA
10	850.4	41.7	200484	2 AC025002	AC025002 Homo sapi
11	839.4	41.1	199863	2 AL354888	AL354888 Homo sapi
12	366	17.9	181838	2 AC106650	AC106650 Homo sapi
13	366	17.9	239888	2 AC093939	AC093939 Rattus no
14	342.6	16.8	355	6 AX336538	AX336538 Sequence
15	273.2	13.4	198772	10 AL606906	AL606906 Mouse DNA
16	249.4	12.2	252	11 G24438	G24438 human STS W
17	224.8	11.0	201861	2 AC103344	AC103344 Rattus no
18	184.8	9.1	199863	2 AL354888	AL354888 Homo sapi
19	160.8	7.9	215711	3 AC014105	AC014105 Drosophila
20	160.8	7.9	215711	3 AC008205	AC008205 Drosophila
21	160.8	7.9	228479	3 AE003749	AE003749 Drosophila
22	131	6.4	77304	2 AC103185	AC103185 Rattus no
23	129.4	6.3	201861	2 AC103344	AC103344 Rattus no
24	126.4	6.2	66680	2 AC100483	AC100483 Mus muscu
25	122.2	6.0	79785	2 AC025068	AC025068 Homo sapi
26	119.8	5.9	77304	2 AC103185	AC103185 Rattus no
27	99	4.9	12242	1 AE013095	AE013095 Thermoplas
28	90.8	4.4	1396	3 AY052773	AY052773 Ctenorhab
29	88.2	4.3	993	8 AB062608	AB062608 Arabidops
30	88.2	4.3	1391	8 AB061403	AB061403 Arabidops
31	88.2	4.3	82411	8 AC068809	AC068809 Genomc s
32	84.6	4.1	1401	8 AB062609	AB062609 Arabidops
33	84.6	4.1	1632	8 AF109376	AF109376 Arabidops
34	84.6	4.1	1634	8 AY080847	AY080847 Arabidops
35	82.6	4.0	1011	8 AB062610	AB062610 Arabidops
36	82.6	4.0	1067	8 AB061401	AB061401 Arabidops
37	82.6	4.0	1401	8 AY125508	AY125508 Arabidops
38	82.6	4.0	83122	2 ATTT20010	ATTT20010 Arabidops
39	82.6	4.0	159290	2 AP004813	AP004813 Oryza sat
40	81.2	4.0	38400	1 SC4H2	SC4H2 Streptomy
41	80.8	3.9	161809	2 AP005173	AP005173 Oryza sat
42	78.8	3.9	14486	1 AE004123	AE004123 Vibrio ch
43	78.8	3.9	167350	2 AC129717	AC129717 Oryza sat
44	78.4	3.8	97005	2 AC113584	AC113584 Tetradon
45	78.2	3.8	193068	8 AP003455	AP003455 Oryza sat

# ALIGNMENTS

Result 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
A98668	A98668	Sequence 3 from Patent WO9910482.	A98668	A98668.1	GI:6781710	unclassified.	unclassified.	1 (bases 1 to 2041).	Lemieux, J. and Hekimi, S.	THE C. ELEGANS GRO-1 GENE	Patent: WO 9910482-A 3 04-MAR-1999;
									LEMIEUX JASON (CA); UNIV MCGILL (CA)		





```

* 40162 40261: gap of unknown length
* 40262 42177: contig of 1916 bp in length
* 42178 42277: gap of unknown length
* 42278 44344: contig of 2067 bp in length
* 44345 44444: gap of unknown length
* 44445 46472: contig of 2028 bp in length
* 46473 46572: gap of unknown length
* 46573 48054: contig of 1482 bp in length
* 48055 48154: gap of unknown length
* 48155 49390: contig of 1236 bp in length
* 49391 49490: gap of unknown length
* 49491 52986: contig of 3496 bp in length
* 52987 53087: gap of unknown length
* 53087 54647: contig of 1561 bp in length
* 54648 54747: gap of unknown length
* 54748 57112: contig of 2365 bp in length
* 57113 58448: contig of 1636 bp in length
* 58449 58949: gap of unknown length
* 58949 60818: contig of 1869 bp in length
* 60818 60917: gap of unknown length
* 60918 63443: contig of 2526 bp in length
* 63444 66294: gap of unknown length
* 66294 66395: gap of unknown length
* 66395 68140: contig of 1746 bp in length
* 68141 68241: gap of unknown length
* 71198 71298: gap of unknown length
* 71299 73990: contig of 2692 bp in length
* 73991 74090: gap of unknown length
* 74091 75872: contig of 1782 bp in length
* 75873 75973: gap of unknown length
* 75973 79070: contig of 3097 bp in length
* 79070 79170: gap of unknown length
* 79170 81463: contig of 2293 bp in length
* 81463 81563: gap of unknown length
* 81563 85047: contig of 3485 bp in length
* 85048 85148: gap of unknown length
* 85148 89223: contig of 4076 bp in length
* 89224 89324: gap of unknown length
* 89324 91525: contig of 2201 bp in length
* 91525 91625: gap of unknown length
* 91625 94768: contig of 3144 bp in length
* 94769 94869: gap of unknown length
* 94869 98303: contig of 3435 bp in length
* 98304 98404: gap of unknown length
* 98404 102480: contig of 4077 bp in length
* 102481 102580: gap of unknown length
* 102581 107197: contig of 4617 bp in length
* 107198 107298: gap of unknown length
* 107298 112447: contig of 5150 bp in length
* 112448 112547: gap of unknown length
* 112548 114968: contig of 2421 bp in length

```

```

Query Match 43.1% Score 38.8; DB 2; Length 181838;
Best Local Similarity 86.0%; Fred. No. 0.0034;
Matches 43; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy 1 TGTGACCTGTGATGATCATCATCTTGGGATGGGATGGGACGCA 50
Db 80342 TGTGACCTTGTGATGATCATCATCTTGGGATGGGACGCA 80391

```

```

RESULT 15
AC093939 239888 bp DNA linear HTG 11-JUL-2002
LOCUS AC093939.4 GI:21723477
DEFINITION Rattus norvegicus clone CH230-172N14, *** SEQUENCING IN PROGRESS
ACCESSION AC093939
VERSION AC093939.4
KEYWORDS HTG; RTGS-PHASE1.
SOURCE Norway rat.

```

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 239888)

AUTHORS

Munzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralung,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkelt,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Darhorne,S.R., David,R., Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Licharge,O., Lien,C., Liu,D., Liu,W., Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,L., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mel,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwoko,S., Ogutu,M., Okwono,G., Oregunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L., Petersen,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Swalek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wellington,S., Williams,G., Williamson,A., Wleczek,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: GEXO  
 Center clone name: CH230-172N14  
 Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 210507 bases at least Q40

LOCUS AC106650 181838 bp DNA linear HTG 13-JUL-2002  
 DEFINITION Rattus norvegicus clone CH230-64124, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 62 unordereded pieces.  
 AC106650  
 AC106650.2 GI:21734730  
 HTG: HTGS\_PHASE1.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 181838)  
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-oshan,F.R., Allen,C.,  
 Aisbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
 Barberia,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabiel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
 Homsa,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichteige,O., Lieu,C., Liu,J., Liu,W., Louisged,H.,  
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogund,M., Okunodu,G.,  
 Ogunyeye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,J.,  
 Rivers,M., Rojas,A., Rojoubokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,  
 Sodergren,E., Sonatke,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,  
 Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G., and Gibbs,R.  
 Direct Submission  
 2 (bases 1 to 181838)  
 Unpublished  
 Direct Submission  
 2 (bases 1 to 181838)  
 Worley,K.C.  
 Direct Submission  
 Submitted (12-JAN-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 181838)  
 Worley,K.C.  
 Direct Submission  
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 11, 2002 this sequence version replaced gi:18139174.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information

Center project name: GLTJ  
 Center clone name: CH230-64124  
 ----- Summary Statistics  
 Sequencing vector: M13  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 115969 bases at least Q40  
 Consensus quality: 123034 bases at least Q30  
 Consensus quality: 127422 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 62 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1  
 1073: contig of 1073 bp in length  
 1074: gap of unknown length  
 1174: contig of 1194 bp in length  
 2368: gap of unknown length  
 2468: contig of 1288 bp in length  
 3756: gap of unknown length  
 3856: contig of 1283 bp in length  
 5139: gap of unknown length  
 5239: contig of 1016 bp in length  
 6255: gap of unknown length  
 6355: contig of 1290 bp in length  
 7645: gap of unknown length  
 7745: contig of 1452 bp in length  
 9197: gap of unknown length  
 9297: contig of 1127 bp in length  
 10424: gap of unknown length  
 10524: contig of 1570 bp in length  
 12094: gap of unknown length  
 12194: contig of 1413 bp in length  
 13607: gap of unknown length  
 13707: contig of 1433 bp in length  
 15140: gap of unknown length  
 15240: contig of 1280 bp in length  
 16520: gap of unknown length  
 16620: contig of 1410 bp in length  
 18030: gap of unknown length  
 18130: contig of 1408 bp in length  
 19538: gap of unknown length  
 19638: contig of 1136 bp in length  
 20774: gap of unknown length  
 20874: contig of 1067 bp in length  
 21941: gap of unknown length  
 22041: contig of 1937 bp in length  
 23977: gap of unknown length  
 24078: contig of 1387 bp in length  
 24078: gap of unknown length  
 25465: contig of 1317 bp in length  
 25565: gap of unknown length  
 26881: contig of 1894 bp in length  
 26982: gap of unknown length  
 28975: contig of 1570 bp in length  
 28976: gap of unknown length  
 30546: contig of 1027 bp in length  
 31673: gap of unknown length  
 31773: contig of 1423 bp in length  
 33196: gap of unknown length  
 33296: contig of 1387 bp in length  
 34783: gap of unknown length  
 34783: contig of 1184 bp in length  
 35966: gap of unknown length  
 36067: contig of 1689 bp in length  
 37756: gap of unknown length  
 37856: contig of 2306 bp in length  
 40161: contig of 2306 bp in length

```
* 28370 28469: gap of 100 bp
* 28470 29153: contig of 684 bp in length
* 28154 29253: gap of 100 bp
* 29254 29971: contig of 718 bp in length
* 29972 30071: gap of 100 bp
* 30072 30770: contig of 699 bp in length
* 30771 30870: gap of 100 bp
* 30871 31581: contig of 711 bp in length
* 31582 31681: gap of 100 bp
* 31682 32380: contig of 699 bp in length
* 32381 32480: gap of 100 bp
* 32481 33187: contig of 707 bp in length
* 33188 33287: gap of 100 bp
* 33288 34012: contig of 725 bp in length
* 34013 34112: gap of 100 bp
* 34113 34831: contig of 719 bp in length
* 34832 34931: gap of 100 bp
* 34932 35627: contig of 696 bp in length
* 35628 35727: gap of 100 bp
* 35728 36446: contig of 719 bp in length
* 36447 36546: gap of 100 bp
* 36547 37245: contig of 699 bp in length
* 37246 37345: gap of 100 bp
* 37346 38077: contig of 732 bp in length
* 38078 38177: gap of 100 bp
* 38178 38879: contig of 702 bp in length
* 38880 38979: gap of 100 bp
* 38980 39688: contig of 709 bp in length
* 39689 39788: gap of 100 bp
* 39789 40500: contig of 712 bp in length
* 40501 40600: gap of 100 bp
* 40601 41317: contig of 717 bp in length
* 41318 41417: gap of 100 bp
* 41418 42116: contig of 699 bp in length
* 42117 42216: gap of 100 bp
* 42217 42923: contig of 707 bp in length
* 42924 43023: gap of 100 bp
* 43024 43739: contig of 716 bp in length
* 43740 43839: gap of 100 bp
* 43840 44570: contig of 731 bp in length
* 44571 44670: gap of 100 bp
* 44671 45381: contig of 711 bp in length
* 45382 45481: gap of 100 bp
* 45482 46191: contig of 710 bp in length
* 46192 46291: gap of 100 bp
* 46292 46986: contig of 695 bp in length
* 46987 47086: gap of 100 bp
* 47087 47789: contig of 703 bp in length
* 47790 47889: gap of 100 bp
* 47890 48590: contig of 701 bp in length
* 48591 48690: gap of 100 bp
* 48691 49414: contig of 724 bp in length
* 49415 49514: gap of 100 bp
* 49515 50220: contig of 706 bp in length
* 50221 50320: gap of 100 bp
* 50321 51010: contig of 690 bp in length
* 51011 51110: gap of 100 bp
* 51111 51823: contig of 713 bp in length
* 51824 51923: gap of 100 bp
* 51924 52648: contig of 725 bp in length
* 52649 52748: gap of 100 bp
* 52749 53466: contig of 718 bp in length
* 53467 53566: gap of 100 bp
* 53567 54295: contig of 729 bp in length
* 54296 54395: gap of 100 bp
* 54396 55105: contig of 710 bp in length
* 55106 55205: gap of 100 bp
* 55206 55935: contig of 730 bp in length
```

```
QY 1 TGTGACCTGTGATGCATTCATTTGGGATCGGATGGGACG 46
Db 25567 TGTGACCTGTGATGCATTCATTTGGGATCGGATGGGACG 25522
```

```
RESULT 13
LOCUS AL606906
DEFINITION AL606906 198772 bp DNA linear ROD 25-JUL-2002
sequence.
ACCESSION AL606906
VERSION AL606906.18 GI:21955487
KEYWORDS HTG.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 198772)
```

```
REFERENCE Direct Submission
AUTHORS Submitted (25-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jul 25, 2002 this sequence version replaced gi:21912638.
```

## COMMENT

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-121J14 is from the RCT-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6.

## FEATURES

## source

1. 198772  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone="RP23-121J14"  
/clone\_1fb="RCT-23"

## BASE COUNT

t

```
Query Match 44.0%; Score 39.6; DB 10; Length 198772;  
Best Local Similarity 91.3%; Pred. No. 0.0017;  
Matches 42; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 TGTGACCTGTGATGCATTCATTTGGGATCGGATGGGACG 46  
|||||  
Db 194238 TGTGACCTGTGATGCATTCATTTGGGATCGGATGGGACG 194283
```

```
RESULT 14  
AC106650
```

\* 137461 140476: contig of 3016 bp in length  
 \* 140477 140576: gap of 100 bp  
 Query Match 52.4% Score 47.2: DB 2: Length 200484:  
 Best Local Similarity 86.7% Pred. No. 2.3e-06:  
 Matches 52: Conservative 0: Mismatches 8: Indels 0: Gaps 0:

OY 31 GATCGCATGAGCGACATTAATCCCAATCCAGCTGACCACTGACAGAGAGAGA 90  
 Db 94041 GATTTCCTTATCAGCGACATTAATCCCAATCCAGCTGACCACTGACAGAGAGAGA 93982

RESULT 12  
 AC025068/c  
 LOCUS  
 DEFINITION Homo sapiens chromosome 1 clone RP11-21702 map 1, LOW-PASS SEQUENCE  
 SAMPLING.  
 AC025068 79785 bp DNA linear HTG 13-JUL-2000  
 AC025068  
 VERSION AC025068.2 GI:7329438  
 KEYWORDS HTG: HTGS\_PHASE0.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 TITLE Homo sapiens chromosome 1, clone RP11-21702  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 79785)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E.,  
 TITLE Homo sapiens chromosome 1, clone RP11-21702  
 JOURNAL Unpublished

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (04-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 330 Charles Street, Cambridge, MA 02141, USA  
 On Mar 26, 2000 this sequence version replaced gi:7158126.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIRB  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L7716  
 Center clone name: 217\_O\_2

\* NOTE: This record contains 98 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.

\* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved. 740: contig of 740 bp in length  
 1  
 741 840: gap of 100 bp  
 841 1543: contig of 703 bp in length  
 1544 1643: gap of 100 bp  
 1644 2350: contig of 707 bp in length  
 2351 2450: gap of 100 bp  
 2451 3158: contig of 708 bp in length  
 3159 3258: gap of 100 bp  
 3259 3978: contig of 720 bp in length  
 3979 4078: gap of 100 bp  
 4079 4775: contig of 698 bp in length  
 4777 4876: gap of 100 bp  
 4877 5593: contig of 717 bp in length  
 5594 5693: gap of 100 bp  
 5694 6412: contig of 719 bp in length  
 6413 6512: gap of 100 bp  
 6513 7238: contig of 726 bp in length  
 7239 7338: gap of 100 bp  
 7339 8055: contig of 717 bp in length  
 8056 8155: gap of 100 bp  
 8156 8870: contig of 715 bp in length  
 8871 8970: gap of 100 bp  
 8971 9691: contig of 721 bp in length  
 9692 9791: gap of 100 bp  
 9792 10493: contig of 702 bp in length  
 10494 10593: gap of 100 bp  
 10594 11326: contig of 733 bp in length  
 11327 11426: gap of 100 bp  
 11427 12137: contig of 711 bp in length  
 12138 12237: gap of 100 bp  
 12238 12950: contig of 713 bp in length  
 12951 13050: gap of 100 bp  
 13051 13751: contig of 701 bp in length  
 13752 13851: gap of 100 bp  
 13852 14552: contig of 701 bp in length  
 14553 14652: gap of 100 bp  
 14653 15372: contig of 720 bp in length  
 15373 15472: gap of 100 bp  
 15473 16183: contig of 711 bp in length  
 16184 16283: gap of 100 bp  
 16284 17005: contig of 722 bp in length  
 17006 17105: gap of 100 bp  
 17106 17825: contig of 720 bp in length  
 17826 17925: gap of 100 bp  
 17926 18638: contig of 714 bp in length  
 18640 18739: gap of 100 bp  
 18740 19457: contig of 718 bp in length  
 19458 19557: gap of 100 bp  
 19558 20268: contig of 712 bp in length  
 20270 20369: gap of 100 bp  
 20370 21069: contig of 700 bp in length  
 21070 21169: gap of 100 bp  
 21170 21895: contig of 726 bp in length  
 21896 21995: gap of 100 bp  
 21996 22708: contig of 714 bp in length  
 22710 22809: gap of 100 bp  
 22810 23507: contig of 698 bp in length  
 23508 23607: gap of 100 bp  
 23608 24297: contig of 690 bp in length  
 24298 24397: gap of 100 bp  
 24398 25107: contig of 710 bp in length  
 25108 25207: gap of 100 bp  
 25208 25936: contig of 729 bp in length  
 25937 26036: gap of 100 bp  
 26037 26749: contig of 713 bp in length  
 26750 26849: gap of 100 bp  
 26850 27562: contig of 713 bp in length  
 27563 27662: gap of 100 bp  
 27663 28369: contig of 707 bp in length

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 81 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved. 1048: contig of 1048 bp in length  
\* 1  
\* 1049 1148: gap of 100 bp  
\* 1149 2470: contig of 1322 bp in length  
\* 2471 2570: gap of 100 bp  
\* 2571 3643: contig of 1073 bp in length  
\* 3644 3743: gap of 100 bp  
\* 3744 4980: contig of 1237 bp in length  
\* 4981 5080: gap of 100 bp  
\* 5081 6352: contig of 1272 bp in length  
\* 6353 6452: gap of 100 bp  
\* 6453 7549: contig of 1097 bp in length  
\* 7550 7649: gap of 100 bp  
\* 7650 8977: contig of 1328 bp in length  
\* 8978 9077: gap of 100 bp  
\* 9078 10103: contig of 1026 bp in length  
\* 10104 10203: gap of 100 bp  
\* 10204 11362: contig of 1159 bp in length  
\* 11363 11462: gap of 100 bp  
\* 11463 12821: contig of 1359 bp in length  
\* 12822 12921: gap of 100 bp  
\* 12922 14007: contig of 1086 bp in length  
\* 14008 14107: gap of 100 bp  
\* 14108 15283: contig of 1176 bp in length  
\* 15284 15383: gap of 100 bp  
\* 15384 16399: contig of 1016 bp in length  
\* 16400 16499: gap of 100 bp  
\* 16500 17804: contig of 1305 bp in length  
\* 17805 17904: gap of 100 bp  
\* 17905 19176: contig of 1272 bp in length  
\* 19177 19276: gap of 100 bp  
\* 19277 20384: contig of 1108 bp in length  
\* 20385 20484: gap of 100 bp  
\* 20485 21689: contig of 1205 bp in length  
\* 21690 21789: gap of 100 bp  
\* 21790 23191: contig of 1402 bp in length  
\* 23192 23291: gap of 100 bp  
\* 23292 24513: contig of 1222 bp in length  
\* 24514 24613: gap of 100 bp  
\* 24614 26135: contig of 1522 bp in length  
\* 26136 26235: gap of 100 bp  
\* 26236 27373: contig of 1138 bp in length  
\* 27374 27473: gap of 100 bp  
\* 27474 28723: contig of 1250 bp in length  
\* 28724 28823: gap of 100 bp  
\* 28824 30217: contig of 1394 bp in length  
\* 30218 30317: gap of 100 bp  
\* 30318 32056: contig of 1739 bp in length  
\* 32057 32156: gap of 100 bp  
\* 32157 33818: contig of 1662 bp in length  
\* 33819 33918: gap of 100 bp  
\* 33919 35156: contig of 1238 bp in length  
\* 35157 35256: gap of 100 bp  
\* 35257 36712: contig of 1456 bp in length  
\* 36713 36812: gap of 100 bp  
\* 36813 38379: contig of 1567 bp in length  
\* 38380 38479: gap of 100 bp  
\* 38480 40190: contig of 1711 bp in length  
\* 40191 40290: gap of 100 bp  
\* 40291 41780: contig of 1490 bp in length  
\* 41781 41880: gap of 100 bp  
\* 41881 44053: contig of 2173 bp in length  
\* 44054 44153: gap of 100 bp  
\* 44154 45474: contig of 1321 bp in length  
\* 45475 45574: gap of 100 bp  
\* 45575 46821: contig of 1247 bp in length

\* 46822 46921: gap of 100 bp  
\* 46922 47302: contig of 381 bp in length  
\* 47303 47402: gap of 100 bp  
\* 47403 49323: contig of 1921 bp in length  
\* 49324 49423: gap of 100 bp  
\* 49424 51140: contig of 1717 bp in length  
\* 51141 51240: gap of 100 bp  
\* 51241 53099: contig of 1859 bp in length  
\* 53100 53199: gap of 100 bp  
\* 53200 54631: contig of 1432 bp in length  
\* 54632 54731: gap of 100 bp  
\* 54732 55751: contig of 1020 bp in length  
\* 55752 55851: gap of 100 bp  
\* 55852 57470: contig of 1619 bp in length  
\* 57471 57570: gap of 100 bp  
\* 57571 60182: contig of 2612 bp in length  
\* 60183 60282: gap of 100 bp  
\* 60283 63065: contig of 2783 bp in length  
\* 63066 63165: gap of 100 bp  
\* 63166 64443: contig of 1278 bp in length  
\* 64444 64543: gap of 100 bp  
\* 64544 66043: contig of 1500 bp in length  
\* 66044 66143: gap of 100 bp  
\* 66144 68623: contig of 2480 bp in length  
\* 68624 68723: gap of 100 bp  
\* 68724 70122: contig of 1399 bp in length  
\* 70123 70222: gap of 100 bp  
\* 70223 72407: contig of 2185 bp in length  
\* 72408 72507: gap of 100 bp  
\* 72508 74635: contig of 2128 bp in length  
\* 74636 74735: gap of 100 bp  
\* 74736 77554: contig of 2819 bp in length  
\* 77555 77654: gap of 100 bp  
\* 77655 80277: contig of 2623 bp in length  
\* 80278 80377: gap of 100 bp  
\* 80378 81798: contig of 1421 bp in length  
\* 81799 81898: gap of 100 bp  
\* 81899 84126: contig of 2228 bp in length  
\* 84127 84226: gap of 100 bp  
\* 84227 86794: contig of 2568 bp in length  
\* 86795 86894: gap of 100 bp  
\* 86895 90311: contig of 3417 bp in length  
\* 90312 90411: gap of 100 bp  
\* 90412 92255: contig of 1844 bp in length  
\* 92256 92355: gap of 100 bp  
\* 92356 94570: contig of 2215 bp in length  
\* 94571 94670: gap of 100 bp  
\* 94671 97163: contig of 2493 bp in length  
\* 97164 97263: gap of 100 bp  
\* 97264 99724: contig of 2461 bp in length  
\* 99725 99824: gap of 100 bp  
\* 99825 103151: contig of 3327 bp in length  
\* 103152 103251: gap of 100 bp  
\* 103252 106370: contig of 3119 bp in length  
\* 106371 106470: gap of 100 bp  
\* 106471 108888: contig of 2418 bp in length  
\* 108889 108988: gap of 100 bp  
\* 108989 111514: contig of 2526 bp in length  
\* 111515 111614: gap of 100 bp  
\* 111615 114360: contig of 2746 bp in length  
\* 114361 114460: gap of 100 bp  
\* 114461 118269: contig of 3809 bp in length  
\* 118270 118369: gap of 100 bp  
\* 118370 122359: contig of 3990 bp in length  
\* 122360 122459: gap of 100 bp  
\* 122460 125669: contig of 3210 bp in length  
\* 125670 125769: gap of 100 bp  
\* 125770 129306: contig of 3537 bp in length  
\* 129307 129406: gap of 100 bp  
\* 129407 132984: contig of 3578 bp in length  
\* 132985 133084: gap of 100 bp  
\* 133085 137360: contig of 4276 bp in length  
\* 137361 137460: gap of 100 bp

```

* 22844 45183: contig of 22340 bp in length
* 45184 45283: gap of 100 bp
* 45284 49581: contig of 4298 bp in length
* 49582 49681: gap of 100 bp
* 49682 58258: contig of 8577 bp in length
* 58259 58358: gap of 100 bp
* 58359 64082: contig of 5724 bp in length
* 64083 64182: gap of 100 bp
* 64183 72764: contig of 8582 bp in length
* 72765 72864: gap of 100 bp
* 72865 86552: contig of 13688 bp in length
* 86553 86652: gap of 100 bp
* 86653 93481: contig of 6829 bp in length
* 93482 93581: gap of 100 bp
* 93582 101614: contig of 8033 bp in length
* 101615 101714: gap of 100 bp
* 101715 107836: contig of 6122 bp in length
* 107837 107936: gap of 100 bp
* 107937 144546: contig of 36610 bp in length
* 144547 144646: gap of 100 bp
* 144647 155892: contig of 11246 bp in length
* 155893 155992: gap of 100 bp
* 155993 167875: contig of 11863 bp in length
* 167876 167975: gap of 100 bp
* 167976 183411: contig of 15436 bp in length
* 183412 183511: gap of 100 bp
* 183512 199863: contig of 16352 bp in length.

```

## FEATURES

## source

```

1. 199863
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="1"
  /clone_1lb="RP11-473p22"
  /clone_1lb="RP11-11.2"
  1. 2278
  /note="assembly:fragment:00765
  fragment_chain:1
  clone_end:17
  vector_side:left"
  2379..22743
  /note="assembly:fragment:01823
  fragment_chain:1"
  22844..45183
  /note="assembly:fragment:01809
  fragment_chain:2"
  45284..49581
  /note="assembly:fragment:01177
  fragment_chain:2"
  49682..58258
  /note="assembly:fragment:00999
  fragment_chain:2"
  58359..64082
  /note="assembly:fragment:01187
  fragment_chain:2"
  64183..72764
  /note="assembly:fragment:00292
  fragment_chain:2"
  72865..86552
  /note="assembly:fragment:00339
  fragment_chain:3"
  86653..93481
  /note="assembly:fragment:01156
  fragment_chain:3"
  93582..101614
  /note="assembly:fragment:01960
  fragment_chain:3"
  101715..107836
  /note="assembly:fragment:00867
  fragment_chain:4"
  107937..144546
  /note="assembly:fragment:01896
  fragment_chain:4"
  144647..155892

```

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

```

/note="assembly:fragment:00307"
155993..167875
/note="assembly:fragment:00528"
167976..183411
/note="assembly:fragment:01170"
183512..199863
/note="assembly:fragment:01894"
53282 a 45915 c 45981 g 53154 t 1531 others

```

```

Query Match 52.48; Score 47.2; DB 2; Length 199863;
Best Local Similarity 86.7%; Pred. No. 2.3e-06;
Matches 52; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

DB 26154 GAATTTGCTTATCAGCGCACATTAATCCAAATCCACTGACCACTGAGAGAGA 90

```

## RESULT 11

## AC025002/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

```

AC025002 200484 bp DNA linear HTG 14-JUN-2000
Homo sapiens chromosome 1 clone RP11-204L3 map 1, *** SEQUENCING IN
PROGRESS *** 81 unordered pieces.
AC025002
AC025002.3 GI:8516086
HTG; HTGS-PHASE1.
Homo sapiens.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 200484)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 1, clone RP11-204L3
Unpublished
2 (bases 1 to 200484)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barra, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choquet, Y., Colangelo, N., Collins, S.,
Collymore, A., Cooke, P., DeArlellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heathford, A., Horton, L.,
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Lacroque, K., Lamazeres, R., Landers, T., Lehotzky, J.,
Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuck, A., McKernan, K., McPheters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,
Pisanil, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rotman, D.,
Roy, A., Santos, R., Schauer, S., Severly, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Yiel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

```

```

TITLE
JOURNAL
COMMENT
Submitted (03-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 14, 2000 this sequence version replaced gi:7770499.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7675
Center clone name: 204_L_3

```

```

/evidence-not_experimental
repeat_region 12351..13402 /note="AluX repeat: matches 19..300 of consensus"
/note="L2 repeat: matches 957..2141 of consensus"
repeat_region 13474..13591 /note="L2 repeat: matches 1468..1747 of consensus"
/note="AluX/AluY repeat: matches 178..295 of consensus"
repeat_region 13596..14185 /note="L2 repeat: matches 2244..2278 of consensus"
/note="L1M3 repeat: matches 7071..7691 of consensus"
repeat_region 14186..14425 /note="L2 repeat: matches 2129..2417 of consensus"
/note="L2 repeat: matches 27..147 of consensus"
repeat_region 14457..14572 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 47..137 of consensus"
repeat_region 15137..15215 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 16285..16510 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 17313..17597 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 20320..20782 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 20785..20932 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 21140..21425 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 21927..22073 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22185..22408 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22409..22704 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22705..22759 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22829..22898 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22932..22989 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 22990..23298 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 23299..23381 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 23401..23576 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 23600..23742 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 24157..24451 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 24618..24684 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 24753..24891 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 25018..25159 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 26215..26275 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 26604..26643 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 26930..27021 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 27380..27733 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 28059..28151 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
misc_feature 28542..30030 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
/evidence-not_experimental
repeat_region 29451..29584 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 30126..30262 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 30283..30460 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 30928..31015 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"
repeat_region 31389..31662 /note="L2 repeat: matches 2278..2701 of consensus"
/note="L2 repeat: matches 2278..2701 of consensus"

Query Match 52.4%, Score 47.2, DB 9, Length 166496;
Best Local Similarity 86.7%, Pred.No. 2.2e-06;
Matches 52; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 31 GATCGGATGAGGCGGACATTAATCCAAATCCACTGACCACTGAAGAAAGA 90
DB 82459 GATTGTGCTTATGACGACATTAATCCAAATCCACTGACCACTGAAGAAAGA 82400

RESULT 10 199863 bp DNA linear HTG 13-JUN-2001
AL354888 Homo sapiens chromosome 1 clone RP11-473P22, *** SEQUENCING IN
LOCUS AL354888/c PROGRESS ***, 16 unordered pieces.
DEFINITION AL354888
VERSION AL354888.6 GI:9863692
KEYWORDS HTG: HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 199863)
AUTHORS Plumb, B.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9213379.

COMMENT ----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
Project Information
Center project name: BA473P22
----- Summary Statistics
Assembly program: XGAP4, version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 189476 bases at least Q40
Consensus quality: 193311 bases at least Q30
Consensus quality: 195510 bases at least Q20
Insert size: 198363; sum-of-contigs
Insert size: 201159; 6.2% error; agarose-fp
Quality coverage: 3.87x in Q20 bases; sum-of-contigs Quality
coverage: 3.98x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2278: contig of 2278 bp in length
* 2279 2378: gap of 100 bp
* 2379 22743: contig of 20365 bp in length
* 22744 22843: gap of 100 bp

```



```

/lab_host="DH10B"
/Note="Vector: PCMV-SPORT6"
406..1386
/product="RIKEN CDNA 2310075G14 gene"
/protein_id="AAH19812.1"
/db_xref="GI:18044186"
/db_xref="locusID:66966"
/translation="MTGKGVDRKVELEKEDCHELHRLISQVDPPEMAKLNHPDRKRV
ASLQVFEETGISHEFLHRAHEGGGLGPRFPNPCIIMLHADQAVLDERLDR
VDDMLAAGLLEELRGFHRNRYLNKINSODYOHGIFOSIGKFEHEVYTTGKCTPE
TSNOLKKGIEALKQVTRKARKONRWNRFLSPGPPVPGYGLVEDVSKWESY
LEPANIYOSFIOGHKPTAMPKYKMYNSENKRSYHMDICDRITIGDEMAHLKSK
SHLHOLKRRRLDPAVSHTSQSNPDCDPRIRIGESSGHNQELKASV"
BASE COUNT      570 a      453 c      541 g      481 t
ORIGIN
Query Match      82.2% Score 74; DB 10; Length 2045;
Best Local Similarity 88.9% Pred. No. 8.9e-17;
Matches 80; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 TGTGACCTGTGTGATGAAATCATTCATTTGGGATCGGAAATGGCAGCGACATAAATCC 60
|||||
Db 1171 TGTGACCTGTGTGACCGGATCATTCATTTGGGATCGGAAATGGCAGCGACATTAATCC 1230
|||||

QY 61 AAATCCACTTGACCACTGAAGAAAGA 90
|||||
Db 1231 AAATCTCACTTGCCCAATGAAGAAAGA 1260
|||||

RESULT 9
HS118J21/c 166496 bp DNA linear PRI 18-FEB-2000
LOCUS HS118J21
DEFINITION Human DNA sequence from clone RPI-118J21 on chromosome 1p34.1-35.3
contains part of the gene for BMP8 (bone morphogenetic protein 8
(osteogenic protein 2)), an 1-myc-proto-oncogene, STS, GSSS and
Cpg Islands.n, complete sequence.
ACCESSION AL033527.26 GI:6456853
VERSION HTG; BMP8; Cpg Island; 1-myc; morphogenetic; oncogene; osteogenic.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 166496)
AUTHORS Ellington, A.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT
Requests: clonerequest@sanger.ac.uk
On Nov 21, 1999 this sequence version replaced gi:6002299.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep RPI-118J21 is
from the library RPI-1 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pcypac2
This sequence is the entire insert of clone RPI-118J21 The true

```

```

FEATURES
source
right end of clone RPI-117L23 is at 38518 in this sequence. This
sequence was generated from part of bacterial clone contigs of
human chromosome 1, constructed by the Sanger Centre Chromosome 1
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1.
location/Qualifiers
1..166496
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="p34.1-35.3"
/clone="RPI-118J21"
/clone_11b="RPI-1"
/complement(830..29381)
/gene="BMP8"
/complement(join(<830..1098,3622..3732,4242..4321,
5153..5347,14838..14986,15379..15568,28682..29381))
/gene="BMP8"
/product="dj118J21.1 (bone morphogenetic protein 8
(osteogenic protein 2))"
/note="match: cDNAs: Em:S77477 Em:X51801 Em:M97017
Em:U39545 Em:J04566 Em:X80992 Em:M60316 Em:M60315
Em:M97016 Em:X56906 Em:AF100787 Em:U40034"
/evidence="not_experimental"
/complement(join(949..1098,3622..3732,4242..4321,
5153..5347,14838..14986,15379..15568,28682..29015))
/gene="BMP8"
/note="match: proteins: Sw:P22004 Sw:P20722 Sw:P22003
Sw:P18075 Tr:O9XG7 Sw:P23359 Sw:P55105 Sw:P49003
Sw:P34820 Sw:P34821"
/codon_start=1
/evidence="not_experimental"
/product="dj118J21.1 (bone morphogenetic protein 8
(osteogenic protein 2))"
/protein_id="CAB75681.1"
/db_xref="GI:7018292"
/translation="MTALPGPLMLGLAICAGGGGPGLRPPGCGPRLGARERDY
OREILVLGLGPRPRAPPAASRLPASRLFLMDLITFMAGDDDGAPAEERLURA
DLVMSFVNVERDRALGHQEPHKEFEREDLTOIPAGEAVTAEFRIYKPSIHLRLT
LHVSMGVVQVDSNRESDLFFLDLQTLRAGDGGWLVDVTAASDCWLRLRKHDIQRL
VYETEDGSHVDVGLAGLIGRAPRSGQPPVFFRASPPIRTPRAVRPLRRRQPKS
NELPOANRLPGIJEIDVHGSHGROVCRHLLYVSFDGLDWTIAQGSAYCECEG
SRPLDSMNATNTHALIQSLVHLMPAPVAKACAPKLSATSVLYYDSSNNVILRHR
NNVAKACGCH"
1655..1748
/note="L2 repeat: matches 1691..1787 of consensus"
1849..2281
/note="L2 repeat: matches 2002..2438 of consensus"
2764..2868
/note="FLAM_C repeat: matches 1..133 of consensus"
/complement(4214..4420)
/gene="BMP8"
/note="match: STS: Em:T63436"
/complement(4252..4438)
/gene="BMP8"
/note="match: STS: Em:AA634039"
5972..6029
/note="MIR repeat: matches 19..75 of consensus"
6293..6588
/note="Aluub repeat: matches 1..307 of consensus"
7787..7873
/note="MIR repeat: matches 57..146 of consensus"
/complement(8369..8876)
/gene="BMP8"
/note="match: GSS: Em:B51067"
9109..9404
/note="Aluub repeat: matches 1..236 of consensus"
9427..9746
/note="Aluub repeat: matches 1..312 of consensus"
10075..10242
/note="match: STS: Em:T41229"
10626..12048
/note="Cpg Island"

```

TITLE Novel nucleic acids and polypeptides  
JOURNAL Patent: WO 022660-A 119 21-MAR-2002;  
HYSEO, INC. (US)

FEATURES  
source  
Location/Qualifiers  
1..2130  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
191..1426  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD34811.1"  
/db\_xref="GI:21438848"  
/translation="MOVYEGLDITINKVSAQORICRHHMISFVDLVTNYTVYDFRN  
RATLIEDIFARDKIPVGGTYIESLAKVYNTKPOEMGETEKYIDRVELEKED  
GLVHKRLSOVDPEMAAKLHPHDKRKVARSIOVEPFGISHSFELHROHREGGGPIG  
GPLKESPCILMLHADQAVLDERLDRVDDMLAGLIEELRDRHRRINQKAVSENSOD  
YOHGIFQSGFKEFEHYLITEGKCTLETSNQLKGLIEALQVTKRARKQNVKRN  
FLSRGPVIVPVYGLEVDVSKWESVLEPALEIVOSFTIOGHKPTATPIKMPYNAENR  
KRSYHLCDCRRIIGDREMAAHIKSKSHLNQLKRRRLDSDAVNTIESQSVSDHNKP  
EPKESGPDNDQELKCSV"

BASE COUNT 620 a 439 c 522 g 549 t

ORIGIN

Query Match 100.0%; Score 90; DB 6; Length 2130;  
Best Local Similarity 100.0%; Pred. No. 8.2e-23;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 60  
|||||  
DB 1211 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 1270  
|||||

OY 61 AATCCCACTTGAACTGAAGAGAAAGA 90  
|||||  
DB 1271 AATCCCACTTGAACTGAAGAGAAAGA 1300  
|||||

RESULT 7  
AF074918 2216 bp RNA 11near PRI 13-DEC-2000  
LOCUS Homo sapiens tRNA isopentenylpyrophosphate transferase precursor  
DEFINITION  
ACCESSION AF074918  
VERSION AF074918.1 GI:11337964  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2216)  
AUTHORS Golovko, A., Hjalms, G., Stibon, F. and Nilander, B.  
TITLE Cloning of a human tRNA isopentenyl transferase  
JOURNAL Gene 258 (1-2), 85-93 (2000)  
MEDLINE 20564178  
PUBMED 11111046  
REFERENCE 2 (bases 1 to 2216)  
AUTHORS Golovko, A. and Hjalms, G.  
TITLE Direct Submission  
JOURNAL Submitted (28-JUN-1998) Dep. of Plant Biology, Swedish University  
of Agricultural Sciences, Box 7080, SE-750 07 uppsala, Sweden  
Location/Qualifiers  
1..2216  
/organism="Homo sapiens"  
/db\_xref="dbEST:AA332152"  
/db\_xref="taxon:9606"  
/note="extension of sequence by 5' RACE"  
1..2216  
/note="incompletely processed mRNA"  
join(15..829,917..1505)  
/note="liphase"  
/codon\_start=1  
/product="tRNA isopentenylpyrophosphate transferase"  
/protein\_id="AA332152.1"  
/db\_xref="GI:11337965"

CDs precursor RNA

/translation="MASVAARAAPVSGSLRGILQRTLPVILGATGKSTLALQGL  
ORLGEIVASDSKQVYEGDITINKVSAQORICRHHMISFVDLVTNYTVDFRRA  
LATIEDIFARDKIPVGGTYIESLAKVYNTKPOEMGETEKYIDRVELEKEDGL  
VLHKRLSOVDPEMAAKLHPHDKRKVARSIOVEPFGISHSFELHROHREGGGPIG  
GPLKESPCILMLHADQAVLDERLDRVDDMLAGLIEELRDRHRRINQKAVSENSOD  
YOHGIFQSGFKEFEHYLITEGKCTLETSNQLKGLIEALQVTKRARKQNVKRN  
FLSRGPVIVPVYGLEVDVSKWESVLEPALEIVOSFTIOGHKPTATPIKMPYNAENR  
KRSYHLCDCRRIIGDREMAAHIKSKSHLNQLKRRRLDSDAVNTIESQSVSDHNKP  
EPKESGPDNDQELKCSV"

BASE COUNT 651 a 461 c 534 g 570 t

ORIGIN

Query Match 100.0%; Score 90; DB 9; Length 2216;  
Best Local Similarity 100.0%; Pred. No. 8.3e-23;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 60  
|||||  
DB 1290 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGGCGACGACATAAAATCC 1349  
|||||

OY 61 AATCCCACTTGAACTGAAGAGAAAGA 90  
|||||  
DB 1350 AATCCCACTTGAACTGAAGAGAAAGA 1379  
|||||

RESULT 8  
BC019812 2045 bp mRNA 11near ROD 07-AUG-2002  
LOCUS Mus musculus, RIKEN cDNA 2310075G14 gene, clone MGC:30541  
DEFINITION IMAGE:5042856, mRNA, complete cds.  
ACCESSION BC019812  
VERSION BC019812.1 GI:18044185  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2045)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (19-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [gcgabs-remail.nih.gov](mailto:gcgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarsson, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Hale, S.M.,  
Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAP Plate: 41 Row: 1 Column: 20.  
Location/Qualifiers  
1..2045  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N"  
/clone="MGC:30541 IMAGE:5042856"  
/tissue\_type="Kidney, normal, 5 month old male mouse."  
/clone\_id="NCI\_CGAP\_Kid14"

```

source
1. 1749
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="COL01371"
/tissue_type="colon"
/clone_lib="COL"
/notes="cloning vector: pME18SFL3"

BASE COUNT      541 a      337 c      408 g      463 t
ORIGIN

Query Match      100.0%; Score 90; DB 9; Length 1749;
Best Local Similarity 100.0%; Pred. No. 8e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGAGCGACATAAATCC 60
      |||
DB 826 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGAGCGACATAAATCC 885
      |||

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
      |||
DB 886 AATCCCACTTGAACCACTGAAGAAAAGA 915
      |||

RESULT 4
BC010741      1844 bp      mRNA      linear      PRI 17-JUL-2001
LOCUS      Homo sapiens, Similar to tRNA isopentenylpyrophosphate transferase,
ACCESSION      BC010741
VERSION      BC010741.1 GI:14789610
KEYWORDS      MGC.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE      1 (bases 1 to 1844)
JOURNAL      Strausberg, R.
      Direct Submission
      Submitted (12-JUL-2001) National Institutes of Health, Mammalian
      Gene Collection (MGC), Cancer Genomics office, National Cancer
      Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
      USA
REMARK      NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT      Contact: MGC help desk
      Email: cgabbs-remail.nih.gov
      Tissue Procurement: ATCC
      CDNA Library Preparation: Life Technologies, Inc.
      CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
      DNA Sequencing by: Sequencing Group at the Stanford Human Genome
      Center, Stanford University School of Medicine, Stanford, CA 94305
      Web site: http://www.shgc.stanford.edu
      Contact: (Dickson, Mark) mcdpaxil.stanford.edu
      Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
      R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAK Prime: 14 Row: P Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA 91: 7019914.
Location/Qualifiers
1. 1844
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:17002 IMAGE:3905836"
/tissue_type="uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
166..1140
/codon_start=-1
/product="Similar to tRNA isopentenylpyrophosphate
transferase"
CDS

```

```

/protein_id="AAH10741.1"
/db_xref="GI:14789611"
/translation="MGTKEVIDRKVELKEDGIVLHKRLSDVPEMAAKLHPDRKY
ARSLQVEEETGISHSEFLRQHTBEGGPIKSNPCLIMLHADDERLDRVD
DMLANGLEELDFHRRYRQKRVNSQDXYGICISIGKEHEYLITREGCITETS
NOLKRGIEALKOVTRKTRAKQNRWKNFSLRPGFTVPYVGLVSDYSKEEVSYLE
PALETVOSEIQGHKPTATPIKMPYNAENKRSYHLCDDRILIGREVAHIKSKSH
LNOLKRRRLRDSDAVNTTIESOSVSPHNKRPKEKSGPQNDDELKCSV"

BASE COUNT      545 a      374 c      454 g      471 t
ORIGIN

Query Match      100.0%; Score 90; DB 9; Length 1844;
Best Local Similarity 100.0%; Pred. No. 8e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGAGCGACATAAATCC 60
      |||
DB 925 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGAGCGACATAAATCC 984
      |||

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
      |||
DB 985 AATCCCACTTGAACCACTGAAGAAAAGA 1014
      |||

RESULT 5
A98668      2041 bp      DNA      linear      PAT 26-JAN-2000
LOCUS      Sequence 3 from Patent WO9910482.
DEFINITION      A98668
ACCESSION      A98668
VERSION      A98668.1 GI:6781710
KEYWORDS      unidentified.
SOURCE      unidentified.
ORGANISM      unidentified.
REFERENCE      1 (bases 1 to 2041)
AUTHORS      Lemieux, J., and Hekimi, S.
TITLE      THE C. ELEGANS GRO-1 GENE
JOURNAL      Patent: WO 9910482-A 3 04-MAR-1999;
      LEMIEUX JASON (CA); UNIV MCGILL (CA)
      Location/Qualifiers
FEATURES
source
1. 2041
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT      589 a      421 c      502 g      529 t
ORIGIN

Query Match      100.0%; Score 90; DB 6; Length 2041;
Best Local Similarity 100.0%; Pred. No. 8.2e-23;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGAGCGACATAAATCC 60
      |||
DB 1121 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGAGCGACATAAATCC 1180
      |||

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
      |||
DB 1181 AATCCCACTTGAACCACTGAAGAAAAGA 1210
      |||

RESULT 6
AX405704      2130 bp      DNA      linear      PAT 14-JUN-2002
LOCUS      Sequence 119 from Patent WO0222260.
DEFINITION      AX405704
ACCESSION      AX405704
VERSION      AX405704.1 GI:21438847
KEYWORDS      human.
SOURCE      human.
ORGANISM      Homo sapiens.
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
      Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
      Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:07:03 ; Search time 278.109 Seconds

(without alignments)  
9418.083 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 90

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAGAAAGA 90

Scoring table: IDENTIFY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sta:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_mu:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sta:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pin:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	1308	9 AY052768	AY052768 Homo sapi
2	90	100.0	1749	9 AK000068	AK000068 Homo sapi
3	90	100.0	1749	9 AK074222	AK074222 Homo sapi
4	90	100.0	1844	9 BC010741	BC010741 Homo sapi
5	90	100.0	2041	6 A98658	A98658 Sequence 3
6	90	100.0	2130	6 AX405704	AX405704 Sequence
7	90	100.0	2216	9 AF074918	AF074918 Homo sapi
8	74	82.2	2045	10 BC019812	BC019812 Mus muscu
9	47.2	52.4	166496	9 HS118721	AL035272 Human DNA
10	47.2	52.4	199863	2 AL354888	AL354888 Homo sapi
11	47.2	52.4	200484	2 AC025002	AC025002 Homo sapi
12	46	51.1	79785	2 AC025068	AC025068 Homo sapi
13	39.6	44.0	198772	10 AL606906	AL606906 Mouse DNA
14	38.8	43.1	181838	2 AC106650	AC106650 Rattus no
15	38.8	43.1	239888	2 AC093939	AC093939 Rattus no
16	38	42.2	201861	2 AC103344	AC103344 Rattus no
17	31.4	34.9	149452	2 AC124912	AC124912 Papio cyn
18	30.8	34.2	61284	2 AC107873	AC107873 Homo sapi
19	30.8	34.2	120007	9 AF064864	AF064864 Homo sapi
20	30.8	34.2	340000	9 HS21C083	AL163283 Homo sapi
21	30.4	33.8	278652	2 AC073809	AC073809 Mus muscu
22	29.4	32.7	118444	2 AC114896	AC114896 Oryza sat
23	29.4	32.7	320250	9 AF117829	AF117829 Homo sapi
24	29	32.2	55227	2 AC100331	AC100331 Mus muscu
25	29	32.2	55227	2 AC100331	AC100331 Mus muscu
26	29	32.2	66821	2 AC121146	AC121146 Mus muscu
27	29	32.2	175390	2 AC114551	AC114551 Mus muscu
28	29	32.2	229957	10 AL670231	AL670231 Mouse DNA
29	28.8	32.0	163115	9 AC105902	AC105902 Homo sapi
30	28.8	32.0	200956	9 AC104165	AC104165 Homo sapi
31	28.6	31.8	152505	2 AC121658	AC121658 Rattus no
32	28.6	31.8	186793	2 AC095657	AC095657 Rattus no
33	28.6	31.8	198839	2 AC094648	AC094648 Rattus no
34	28.4	31.6	167439	2 AC130273	AC130273 Papio cyn
35	28.4	31.6	176552	2 AC130272	AC130272 Papio cyn
36	28.4	31.6	190352	2 AC101735	AC101735 Mus muscu
37	28.4	31.6	210583	2 AC117723	AC117723 Mus muscu
38	28.2	31.3	41633	9 AC096585	AC096585 Homo sapi
39	28.2	31.3	74531	2 AC015829	AC015829 Homo sapi
40	28.2	31.3	97255	9 AL590482	AL590482 Human DNA
41	28.2	31.3	114041	2 AP001866	AP001866 Homo sapi
42	28.2	31.3	157091	2 AC025484	AC025484 Homo sapi
43	28.2	31.3	162289	4 AC097230	AC097230 Sus scrofa
44	28.2	31.3	163460	2 AC069378	AC069378 Homo sapi
45	28.2	31.3	166039	2 AC044855	AC044855 Homo sapi

#### ALIGNMENTS

RESULT 1  
LOCUS AY052768 1308 bp mRNA linear PRI 17-OCT-2001  
DEFINITION Homo sapiens tRNA isopentenyl transferase mRNA, partial cds.  
ACCESSION AY052768  
VERSION AY052768.1 GI:16209578  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1308)  
Lemieux,J., Lakowski,B., Webb,A., Meng,Y., Ubach,A., Bussiere,F., Barnes,T. and Reklms.  
TITLE Regulation of physiological rates in Caenorhabditis elegans by a

Pred. No. is the number of results predicted by chance to have a

Db 28 GAGGGCTCAGAGAGCTTATATCATGGCAGAGAGGCGAAGGGAAGCAAGCAACCTTCTTCA 87

Qy 62 AATCCCACTTGAACCACTGAAGAAAAGA 90

Db 88 CATGGCAGCAGGAGAGAAAGAGAAAGGA 116

Search completed: April 21, 2003, 13:37:15  
Job time : 30.58 secs









PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226682.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241121.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251889.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

## (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI: 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis.

Disclosure; SEQ ID NO 31146; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/hematopoietic-related diseases, especially cancers and cancer metastases of hematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/hematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.

Sequence 3616 BP; 1000 A; 855 C; 784 G; 977 T; 0 other;

Query Match 30.2%; Score 27.2; DB 22; Length 3616;  
Best Local Similarity 58.8%; Pred. No. 6.7; Indels 0;  
Matches 47; Conservative 0; Mismatches 33; Gaps 0;

XX The invention relates to antisense inhibitors of genes essential to  
CC prokaryotic cellular proliferation, their use in identifying the  
CC genes, their use in the discovery of novel antibiotics, the essential  
CC genes themselves and the encoded proteins. The prokaryotes used are  
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
CC invention is also useful for the identification of potential new targets  
CC for antibiotic development. The antisense nucleic acids can also be used  
CC to identify proteins used in proliferation, to express these proteins,  
CC and to obtain antibodies capable of binding to the expressed proteins.  
CC The proteins can be used to screen compounds in rational drug discovery  
CC programmes. The antisense nucleic acid sequence is also useful to screen  
CC for homologous nucleic acids which are required for cell proliferation in  
CC a wide variety of organisms. The present sequence encodes an  
CC essential prokaryotic cellular proliferation protein.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2703 BP; 895 A; 566 C; 651 G; 591 T; 0 other;  
Query Match 30.4%; Score 27.4; DB 23; Length 2703;  
Best Local Similarity 57.6%; Pred. No. 5.2;  
Matches 49; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
QY 2 GTGACCTCTGTGATGCATCATTCATGGGATCGCGAATGGCGAGCGACATAAATCCA 61  
DB 2525 GTGTTATCGGTATGCGATGTCGTTATCTATGATGGCAACTCGCAACTTGAACACTACA 2584  
QY 62 AATCCACTTGAACCACTGAGCAA 86  
DB 2585 AAGACGACGGAAGAAGTGACAAA 2609  
RESULT 8  
AB17248/c  
ID AB17248 standard; DNA; 6399 BP.  
XX  
AC AB17248;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 3217.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
XX  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PMD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX  
PS Claim 1; SEQ ID NO 3217; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins  
CC (AB557737-AB572072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 6399 BP; 1596 A; 1565 C; 1565 G; 1673 T; 0 other;  
Query Match 30.4%; Score 27.4; DB 23; Length 6399;  
Best Local Similarity 62.3%; Pred. No. 6.6;  
Matches 43; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
QY 1 TGTGACCTCTGTGATGCATCATTCATGGGATCGCGAATGGCGAGCGACATAAATCC 60  
DB 3921 TGTCAATATGCGAACGCGATTTCGTTGGGAGTACCAATGGGACTGATATGAAGTCC 3862  
QY 61 AATCCAC 69  
DB 3861 AACAAACAC 3853  
RESULT 9  
AAK76334  
ID AAK76334 standard; DNA; 3616 BP.  
XX  
AC AAK76334;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31146.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX  
KW cytosolic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
XX  
PR 04-FEB-2000; 2000US-0180628.  
XX  
PR 24-FEB-2000; 2000US-0184664.  
XX  
PR 02-MAR-2000; 2000US-0186350.  
XX  
PR 16-MAR-2000; 2000US-0189874.  
XX  
PR 17-MAR-2000; 2000US-0190076.  
XX  
PR 18-APR-2000; 2000US-0198123.  
XX  
PR 19-MAY-2000; 2000US-0205515.  
XX  
PR 07-JUN-2000; 2000US-0209467.  
XX  
PR 28-JUN-2000; 2000US-0214886.  
XX  
PR 30-JUN-2000; 2000US-0215135.  
XX  
PR 07-JUL-2000; 2000US-0216647.  
XX  
PR 07-JUL-2000; 2000US-0216880.  
XX  
PR 11-JUL-2000; 2000US-0217487.  
XX  
PR 11-JUL-2000; 2000US-0217496.  
XX  
PR 14-JUL-2000; 2000US-0218290.  
XX  
PR 26-JUL-2000; 2000US-0220963.  
XX  
PR 26-JUL-2000; 2000US-0220964.  
XX  
PR 14-AUG-2000; 2000US-0224518.  
XX  
PR 14-AUG-2000; 2000US-0224519.  
XX  
PR 14-AUG-2000; 2000US-0225213.  
XX  
PR 14-AUG-2000; 2000US-0225214.  
XX  
PR 14-AUG-2000; 2000US-0225266.  
XX  
PR 14-AUG-2000; 2000US-0225267.  
XX  
PR 14-AUG-2000; 2000US-0225268.  
XX  
PR 14-AUG-2000; 2000US-0225270.

(II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 1776 BP; 491 A; 450 C; 464 G; 371 T; 0 other;

Query Match 30.7%; Score 27.6; DB 23; Length 1776;  
Best Local Similarity 60.8%; Pred. No. 3.9;  
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 TGTGACCTGTGTATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACATAAAATCC 60  
|||||  
Db 117 TGTGACCTGTGTATCGAATCATCATTTGGGATCGCGAATGGCGAGCGAATAATCTCTCC 176

QY 61 AATATCCCACTTGAA 74  
| | | | |  
Db 177 TCAACACACTGCA 190

RESULT 6  
AAS67229  
ID AAS67229 standard; cDNA; 1780 BP.

AC AAS67229;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #3033.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HSE-) HXSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PSDB: ABG03042.

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

PS Claim 1; SEQ ID No 3033; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human CC diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 1780 BP; 491 A; 452 C; 465 G; 372 T; 0 other;

Query Match 30.7%; Score 27.6; DB 23; Length 1780;  
Best Local Similarity 60.8%; Pred. No. 3.9;  
Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1 TGTGACCTGTGTATCGAATCATCATTTGGGATCGCGAATGGCGAGCGACATAAAATCC 60  
|||||  
Db 117 TGTGACCTGTGTATCGAATCATCATTTGGGATCGCGAATGGCGAGCGAATAATCTCTCC 176

QY 61 AATATCCCACTTGAA 74  
| | | | |  
Db 177 TCAACACACTGCA 190

RESULT 7  
AAS55928  
ID AAS55928 standard; DNA; 2703 BP.

AC AAS55928;

DT 13-FEB-2002 (first entry)

DE Streptococcus pneumoniae DNA for cellular proliferation protein #499.

DE Antisense; ds; prokaryotic cellular proliferation gene;

KW antibiotic; antibacterial; drug design.

OS

PN WO200170955-A2.

PD 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207227P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

PA (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

DR WPI: 2001-611495/70.

DR P-PSDB: AAU38069.

PT New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids -

PS Claim 27; Seq ID No 9565; 511pp; English.

CC genes correlated with a cancerous state of a mammalian cell, comprising  
CC detecting at least one differentially expressed gene product in a test  
CC sample derived from a cell suspected of being cancerous, where detection  
CC of the differentially expressed gene product is correlated with a  
CC cancerous state of the cell from which the test sample was detected.  
CC The polynucleotide sequences can be used in a method for detecting  
CC differentially expressed genes correlated with a cancerous state of a  
CC mammalian cell. The polynucleotides can also be used as probes for  
CC detecting and mapping related genes. They can be used in diagnosis and  
CC prognosis of diseases and disorders (e.g. identification of  
CC pre-metastatic or metastatic cancerous states, stages of cancer, or  
CC responsiveness of cancer to therapy). This is particularly for breast  
CC cancer, oestrogen receptor-positive breast cancer, oestrogen receptor-  
CC negative breast cancer, lung cancer, and colon cancer.  
XX

XX Sequence 300 BP; 83 A; 69 C; 72 G; 76 T; 0 other;

Query Match 48.9%; Score 44; DB 21; Length 300;  
Best Local Similarity 100.0%; Pred. No. 2.4e-06;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTGGGGATCGCGAATGGGC 44  
|||||

Db 257 TGTGACCTCTGTGATCGAATCATTCATTGGGGATCGCGAATGGGC 300

#### RESULT 4

AAS67227  
ID AAS67227 standard; cDNA; 1248 BP.

XX AAS67227;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #3031.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG03040.

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity

XX Claim 1; SEQ ID NO 3031; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

XX Sequence 1248 BP; 334 A; 343 C; 289 G; 282 T; 0 other;

Query Match 30.7%; Score 27.6; DB 23; Length 1248;  
Best Local Similarity 60.8%; Pred. No. 3.6;

Matches 45; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTGGGGATCGCGAATGGGC 60  
|||||

Db 423 TGTGACCTCTGTGATCGAATCATTCATTGGGGATCGCGAATGGGC 482

OY 61 AATCCGACTTGA 74  
|||||

Db 483 TCAACACACTGCAA 496

#### RESULT 5

AAS64323  
ID AAS64323 standard; cDNA; 1776 BP.

XX AAS64323;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #127.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG00136.

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity

XX Claim 1; SEQ ID NO 127; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving

PT useful for enhancing longevity of a host and inhibiting tumour  
 PT formation  
 XX  
 XX Claim 8; Fig 8; 93pp; English.  
 XX  
 CC The present sequence represents the human homologue of gro-1, and  
 CC is referred to as hgro-1. The specification describes the five genes  
 CC of the Caenorhabditis elegans gro-1 operon (AA36071). The operon  
 CC contains the gro-1 gene (AA36072), the gop-1 gene (AA36074), the gop-2  
 CC gene (AA36075), the gop-2 gene (AA36075), and the hap-1 gene  
 CC (AA36077). The gro-1 gene can be used in a method for the diagnosis  
 CC and/or prognosis of cancer in a patient. Transgenic mice containing a  
 CC gene knock-out of a murine gene homologue of the gro-1 gene are useful  
 CC as models of aging and cancer. The proteins encoded by the genes are  
 CC useful for identifying compounds that affect the enzymatic activity  
 CC of these proteins. In order to enhance longevity of a host and inhibit  
 CC tumour formation. The gro-1 gene, together with the gop-1, gop-2,  
 CC gop-3 and hap-1 genes enables study of a physiological clock.  
 CC  
 SQ Sequence 2041 BP; 589 A; 421 C; 502 G; 529 T; 0 other;  
 QY  
 Query Match 100.0%; Score 90; DB 20; Length 2041;  
 Best Local Similarity 100.0%; Pred. No. 6e-23;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1121 TGTGACCTCTGTGATCGAATCATTCATGGGATCGCAATGGCGCACATATAATCC 60  
 1121 TGTGACCTCTGTGATCGAATCATTCATGGGATCGCAATGGCGCACATATAATCC 1180  
 QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
 1181 AATCCCACTTGAACCACTGAAGAAAAGA 1210  
 Db 1181 AATCCCACTTGAACCACTGAAGAAAAGA 1210  
 QY  
 RESULT 2  
 ABRN59708  
 ID ABRN59708 standard; cDNA; 2130 BP.  
 XX  
 AC ABRN59708;  
 XX  
 XX 28-JUN-2002 (first entry)  
 DT  
 XX  
 XX Novel human coding sequence SEQ ID NO: 119.  
 DE  
 XX  
 XX Human: anti-naemic; vulnerary; anti-inflammatory; immunomodulator;  
 KM anti-infectivity; cerebroprotective; cytoskeletal; rheumatic; gene therapy;  
 KM neuroprotective; antiParkinsonian; protein therapy; EST;  
 KM expressed sequence tag; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200222660-A2.  
 XX  
 PD 21-MAR-2002.  
 XX  
 PF 10-SEP-2001; 2001WO-US26015.  
 XX  
 PR 11-SEP-2000; 2000US-0659671.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F,  
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 XX  
 DR WPI; 2002-292408/33.  
 XX  
 DR P-PSDB; ABB97295.  
 XX  
 PT An isolated polynucleotide for treating diseases associated with its  
 PT encoded polypeptide such as cancer and multiple sclerosis -  
 XX  
 PS Claim 1; SEQ ID NO 119; 509pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of 444

CC novel human proteins. These were isolated from expressed sequences tags  
 CC (ESTs). They can be used to stimulate cell growth, to regulate  
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
 CC Parkinson's disease. The present sequence is a coding sequence of the  
 CC invention.  
 CC  
 SQ Sequence 2130 BP; 620 A; 439 C; 522 G; 549 T; 0 other;  
 QY  
 Query Match 100.0%; Score 90; DB 24; Length 2130;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-23;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1211 TGTGACCTCTGTGATCGAATCATTCATGGGATCGCAATGGCGCACATATAATCC 1270  
 1211 TGTGACCTCTGTGATCGAATCATTCATGGGATCGCAATGGCGCACATATAATCC 1270  
 QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
 1271 AATCCCACTTGAACCACTGAAGAAAAGA 1300  
 Db 1271 AATCCCACTTGAACCACTGAAGAAAAGA 1300  
 QY  
 RESULT 3  
 AAA00911  
 ID AAA00911 standard; cDNA; 300 BP.  
 XX  
 AC AAA00911;  
 XX  
 XX 19-MAY-2000 (first entry)  
 DT  
 XX  
 XX Human colon cancer cell line polynucleotide sequence SEQ ID NO:902.  
 DE  
 XX  
 XX Human: colon cancer; tumour; diagnosis; gene expression product;  
 KM probe; detection; cancerous state; metastasis; identification;  
 KM breast cancer; oestrogen receptor-positive breast cancer; therapy;  
 KM oestrogen receptor-negative breast cancer; lung cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO958675-A2.  
 XX  
 PD 18-NOV-1999.  
 XX  
 PF 13-MAY-1999; 99WO-US10602.  
 XX  
 PR 14-MAY-1998; 98US-0085426.  
 XX  
 PR 15-MAY-1998; 98US-0085537.  
 XX  
 PR 15-MAY-1998; 98US-0085696.  
 XX  
 PR 21-OCT-1998; 98US-0105234.  
 XX  
 PR 27-OCT-1998; 98US-0105877.  
 XX  
 PA (CHIR) CHIRON CORP.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
 PI Reinhard C, Glese K, Randazzo F, Kennedy GC, Pot D, Kassam A;  
 PI Lamson G, Drmanac R, Crivenjakov R, Dickson M, Drmanac S, Labat I;  
 PI Leshkowitz D, Kila D, Garcia V, Jones LW, Stache-Crain B;  
 XX  
 DR WPI; 2000-126369/11.  
 XX  
 DR Polynucleotide library used to determine cancerous states of mammalian  
 PT cells -  
 PT  
 XX  
 XX Claim 1; Page 399; 1097pp; English.  
 CC  
 CC AAA00010 to AAA02716 represent polynucleotides isolated from cDNA  
 CC libraries constructed from human colon cancer cell lines. The present  
 CC invention also describes a method of detecting differentially expressed

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:03:18 ; Search time 24.58 Seconds

(without alignments)  
8245.719 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 90  
Sequence: 1 TCGACCTCTGTGATCGAAT.....TGACCACTGAAGAAAGA 90

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008  
Listing first 45 summaries

Database :

N\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	100.0	2041	20	AAK36073	Human homologue of
2	90	100.0	2130	24	ABN59708	Human homologue of
3	44	48.9	300	21	AAA00911	Human colon cancer
4	27.6	30.7	1248	23	AAK67227	DNA encoding novel
5	27.6	30.7	1776	23	AAK64323	DNA encoding novel
6	27.6	30.7	1780	23	AAK67229	DNA encoding novel
7	27.4	30.4	2703	23	AAK55928	Streptococcus pneu
8	27.4	30.4	6399	23	ABL17248	Drosophila melanog
9	27.2	30.2	3616	22	AAK76334	Human immune/haema

10	26.8	29.8	20303	18	AAT71699	Human deoxyctidyl
11	26.8	29.8	26764	18	AAT71696	Human deoxyctidyl
12	26.6	29.6	255	19	AAK11458	Human blallelic po
13	26.6	29.6	255	19	AAK11457	Human blallelic po
14	26.6	29.6	344	20	AAH87247	Human single nucle
15	26.6	29.6	344	20	AAH87261	Human single nucle
16	26.2	29.1	3364	23	ABL05458	Drosophila melanog
17	26.2	29.1	5299	23	ABL07890	Drosophila melanog
18	26	28.9	1472	23	AAK75217	DNA encoding novel
19	26	28.9	2640	22	AAK52965	Human endocrine po
20	26	28.9	3423	22	AAK58249	Human polynucleoti
21	26	28.9	3443	20	AAK08694	Novel nucleotide s
22	26	28.9	3747	23	ABL12801	Drosophila melanog
23	26	28.9	4346	22	AAK67556	Human cardiovascu
24	26	28.9	11116	22	AAK59270	Human polynucleoti
25	25.8	28.7	861	19	AAK43021	Streptococcus pneu
26	25.8	28.7	2793	21	AAK09487	Streptococcus pneu
27	25.4	28.2	1815	20	AAK60607	Polynucleotide seq
28	25.4	28.2	1815	20	AAK60608	Polynucleotide seq
29	25.4	28.2	1815	22	AAK00567	Truncated cryt to
30	25.4	28.2	1815	22	AAK00567	Truncated cryt to
31	25.4	28.2	7028	23	ABL30122	Drosophila melanog
32	25.4	28.2	11313	23	ABL12356	Drosophila melanog
33	25.4	28.2	20776	23	ABL30040	Drosophila melanog
34	25.2	28.0	90	24	ABK36689	Human DNA encoding
35	25.2	28.0	408	23	AAK88099	Human DNA encoding
36	25.2	28.0	471	20	AAK32049	Human MTH2 relate
37	25.2	28.0	471	22	AAK90306	HEONN73R cDNA clon
38	25.2	28.0	502	24	AAK38928	Exon 1 of Siclec8-
39	25.2	28.0	626	21	AAK43007	Human 5' EST isoia
40	25.2	28.0	1090	22	AAK94448	Human foetal cDNA,
41	25.2	28.0	1554	24	ABK43364	DNA encoding stail
42	25.2	28.0	1592	24	AAK38935	Siclec8-L coding c
43	25.2	28.0	1597	20	AAK17773	SAF-2 polypeptide
44	25.2	28.0	2898	22	AAK22575	Human cDNA encodin
45	25.2	28.0	2898	22	AAK22811	Human cDNA encodin

#### ALIGNMENTS

RESULT 1  
ID AAK36073 standard; DNA: 2041 BP.  
XX AAK36073:  
AC  
XX  
XX 15-JUL-1999 (first entry)  
DE Human homologue of gro-1, referred to as hgro-1.  
XX  
XX gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene;  
KW hap-1 gene; cancer; aging; longevity; tumour formation;  
KW physiological clock; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W09910482-A1.  
PN  
XX  
XX 04-MAR-1999.  
PD  
XX  
XX 20-AUG-1998; 98WO-CA00803.  
PF  
XX  
XX 25-AUG-1997; 97CA-2210251.  
PR  
XX  
XX (UYMC-) UNIV MCGILL.  
PA  
XX  
XX Barnes T, Hekimi S, Lakowski B, Lemieux J;  
PI WPI; 1999-190615/16.  
XX  
XX Molecular identity of the gro-1 gene - useful for cancer diagnosis  
PT and/or prognosis, and where compounds affecting encoded proteins are

Search completed: April 21, 2003, 16:30:12  
Job time : 12.3078 secs



```

: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 210:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 11378 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-527-210

Query Match 26.9%; Score 24.2; DB 4; Length 11378;
Best Local Similarity 55.3%; Pred. No. 12;
Matches 47; Conservative 0; Mismatches 39; Indels 0; Gaps 0

QY 2 GTGACCTGTGTGATCGAATCATCATTTGGGATGCGCAATGGCGACCGACATPAAATCCA 61
Db 6393 GTGTATCCGATGATGGTGTCTTATCTATGATGATGATGTAAGTCCAGACTCGAAGCTTGAAGAACTATA 6452
:
:
QY 62 AATCCCACTGTGACCACTGAAGAA 86
Db 6453 AAGACGACGTGAAGAAAGTGACAAA 6477
:
:
RESULT 12
: US-08-484-891-7/c
: Sequence 7, Application US/08484891
: Patent No. 5935935
: GENERAL INFORMATION:
: APPLICANT: Connolly, Sheila
: APPLICANT: Kaleko, Michael
: APPLICANT: Smith, Theodore
: TITLE OF INVENTION: Adenoviral Vectors for
: TITLE OF INVENTION: Treatment of Hemophilia
: NUMBER OF SEQUENCES: 7
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Carella, Byrne, Bain, Gillfillan,
: ADDRESSEE: Cecchi, Stewart & Olstein
: STREET: 6 Becker Farm Road
: CITY: Roseland
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07068
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch diskette
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: Wordperfect 5.1
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,891
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/218,335
: FILING DATE: 25-MAR-1994
: APPLICATION NUMBER: 08/074,920
: FILING DATE: 10-JUN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Olstein, Elliot M.
: REGISTRATION NUMBER: 24,025
: REFERENCE/DOCKET NUMBER: 271010-273
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-994-1700
: TELEFAX: 201-994-1744
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4629 bases
: TYPE: nucleic acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA primer
FEATURE:
NAME/KEY: Factor VIIl cDNA with
NAME/KEY: B domain deleted
US-08-484-891-7

Query Match      26.7%; Score 24; DB 2; Length 4629;
Best Local Similarity 58.3%; Pred. No. 10;
Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY    18 AATCATCTGGGAGTCGAATGGCGACGCACATAAATCCAAATCCACTTGACCA 77
Db    2446 AATCTTCTTTCTTCACTTTCACTGATATGATATCATCATCAGTAATTCTTGTATCTG 2387
          ||||| | | | | | | | | | | | | | | | | | | | | | | |
          ||||||| | |

DB    2386 ACTGAGAAGTAG 2375

RESULT 13
US-08-717-294-41/C
Sequence 41, Application US/08717294
Patent No. 6114148
GENERAL INFORMATION:
APPLICANT: SEED, BRIAN
APPLICANT: HAAS, JURGEN
TITLE OF INVENTION: HIGH LEVEL EXPRESSION OF
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 110
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Ebling LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASCSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/717,294
FILING DATE: 20-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Ebling, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/345001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO.: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 4670 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-717-294-41

Query Match      26.7%; Score 24; DB 3; Length 4670;
Best Local Similarity 58.3%; Pred. No. 10;
Matches 42; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

OY    18 AATCATCTGGGAGTCGAATGGCGACGCACATAAATCCAAATCCACTTGACCA 77

```



GENERAL INFORMATION:  
APPLICANT: Zalacain, Magdalena  
APPLICANT: Burnham, Martin K. R.  
APPLICANT: Biswas, Sanjoy  
APPLICANT: Brown, James R.  
APPLICANT: Ingraham, Karen A.  
APPLICANT: Chalker, Allison F.  
APPLICANT: So, Chi Young  
APPLICANT: Holmes, David J.  
APPLICANT: Trautl, Christopher M.  
APPLICANT: Warren, Richard L.  
APPLICANT: Mathie, Thomas B.  
TITLE OF INVENTION: Infb  
FILE REFERENCE: GM10175  
CURRENT APPLICATION NUMBER: US/09/574,912  
CURRENT FILING DATE: 2000-05-19  
PRIOR APPLICATION NUMBER: 09/283,763  
PRIOR FILING DATE: 1999-04-01  
PRIOR APPLICATION NUMBER: 60/105,985  
PRIOR FILING DATE: 1999-10-28  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 2793  
TYPE: DNA  
ORGANISM: Streptococcus pneumoniae  
US-09-574-912-1

Query Match  
Best Local Similarity 28.7%; Score 25.8; DB 4; Length 2793;  
Matches 48; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 2 GTGACCTGTGATGATCATCATTTGGGATCGGATGGCGACGACATTAATCA 61  
DB 2615 GTGTATCGGTATGCTGTCTTATCTATGATGGCGAAGCTGTAACACTACA 2674

QY 62 AATCCACTGTGAACCACTGAGAA 86  
DB 2675 AAGATGACGTGAAGAGTGACAA 2699

RESULT 6  
US-09-178-252-7/c  
Sequence 7, Application US/09178252  
Patent No. 6218188  
GENERAL INFORMATION:  
APPLICANT: Cardineau, Guy A.  
APPLICANT: Steilman, Steven J.  
APPLICANT: Narva, Kenneth E.  
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins  
FILE REFERENCE: MA-714XC2  
CURRENT APPLICATION NUMBER: US/09/178,252  
CURRENT FILING DATE: 1998-10-23  
EARLIER APPLICATION NUMBER: 60/065,215  
EARLIER FILING DATE: 1997-11-12  
EARLIER APPLICATION NUMBER: 60/076,445  
EARLIER FILING DATE: 1998-03-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 7  
LENGTH: 1815  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic B.t. toxin gene  
US-09-178-252-7

Query Match  
Best Local Similarity 28.2%; Score 25.4; DB 4; Length 1815;  
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 3 TGACCTGTGATGATCATCATTTGGGATCGGATGGCGACGACAT 53  
DB 430 TGATGGCGGTGATCAAGCGCTCATCGTGTGCGAAGCGATACGACAT 380

DB 430 TGATGGCGGTGATCAAGCGCTCATCGTGTGCGAAGCGATACGACAT 380  
RESULT 7  
US-09-178-252-8/c  
Sequence 8, Application US/09178252  
Patent No. 6218188  
GENERAL INFORMATION:  
APPLICANT: Cardineau, Guy A.  
APPLICANT: Steilman, Steven J.  
APPLICANT: Narva, Kenneth E.  
TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins  
FILE REFERENCE: MA-714XC2  
CURRENT APPLICATION NUMBER: US/09/178,252  
CURRENT FILING DATE: 1998-10-23  
EARLIER APPLICATION NUMBER: 60/065,215  
EARLIER FILING DATE: 1997-11-12  
EARLIER APPLICATION NUMBER: 60/076,445  
EARLIER FILING DATE: 1998-03-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 8  
LENGTH: 1815  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic B.t. toxin gene  
US-09-178-252-8

Query Match  
Best Local Similarity 28.2%; Score 25.4; DB 4; Length 1815;  
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 3 TGACCTGTGATGATCATCATTTGGGATCGGATGGCGACGACAT 53  
DB 430 TGATGGCGGTGATCAAGCGCTCATCGTGTGCGAAGCGATACGACAT 380

RESULT 8  
US-09-038-832-3/c  
Sequence 3, Application US/09038832  
Patent No. 6146845  
GENERAL INFORMATION:  
APPLICANT: KIKLY, KRISTINE  
APPLICANT: ERICKSON-MILLER, CONNIE  
TITLE OF INVENTION: Sialoadhesin Family Member-2  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RATNER & PRESTIA  
STREET: P.O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/038,832  
FILING DATE: 11-MAR-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/041,886  
FILING DATE: 02-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-50018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700





```
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1128
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1128
```

```
Query Match          27.6%; Score 24.8; DB 10; Length 653;
Best Local Similarity 63.3%; Pred. No. 14;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
```

```
QY      4  GACCTCTGTGATCATCATTTGGGATCGCGAATGGGCGACGACATAAATCCAAA 63
DB      268 GACCTTAGAGACCTTATATCATGTGTGAGAGCGAAGGGGAGACAAACAGAGAGAGAAA 327
```

## RESULT 15

```
US-09-925-302-212
; Sequence 212, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 212
; LENGTH: 1529
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-212
```

```
Query Match          27.6%; Score 24.8; DB 10; Length 1529;
Best Local Similarity 56.0%; Pred. No. 20;
Matches 47; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
```

```
QY      7  CTCTGTGATCGAATCATCATTTGGGATCGCGAATGGGCGACGACATAAATCCAAATCC 66
DB      403 CTCTGCTCTCCCAAGCCCTTCTCATGTGTGGCGAGTGGAAGGACATATAAGGAATCC 462
```

```
QY      67  CACTTGAAACCACTGAAGAAAGA 90
DB      463 CATTGGAAGTACAGAAAGAAAGA 486
```

Search completed: April 21, 2003, 18:47:45  
Job time : 13.3895 secs

RESULT 12  
US-09-972-714-9/c

RESULT 14  
 US-09-764-860-1128  
 ; Sequence 1128, Application US/0974860.  
 ; Patent No. US20020094953A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kossen et al.,  
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

Query Match 28.9% Score 26; DB 9; Length 4346;  
Best Local Similarity 76.2%; Pred. No. 10;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATAAATCCAAATCCACTTGAAACCACTGAGAGAAAG 89  
DB 4138 GAACCTAAATCTGCTGCCACTGAGCAGTGAAGAGAGG 4179

RESULT 6  
US-09-764-869-2256

; Sequence 2256, Application US/09764869  
; Patent No. US20020061521A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC007

; CURRENT APPLICATION NUMBER: US/09/764,869

; PRIOR FILING DATE: 2001-01-17

; PRIOR APPLICATION DATA REMOVED - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 2442

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 2256

; LENGTH: 4346

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-869-2256

Query Match 28.9% Score 26; DB 10; Length 4346;  
Best Local Similarity 76.2%; Pred. No. 10;  
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATAAATCCAAATCCACTTGAAACCACTGAGAGAAAG 89  
DB 4138 GAACCTAAATCTGCTGCCACTGAGCAGTGAAGAGAGG 4179

RESULT 7

US-09-826-660-7/c

; Sequence 7, Application US/09826660

; Patent No. US20010026940A1

; GENERAL INFORMATION:

; APPLICANT: Cardineau, Guy A.

; APPLICANT: Steilman, Steven J.

; APPLICANT: Narva, Kenneth E.

; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins

; FILE REFERENCE: NA-714XC2D1

; CURRENT APPLICATION NUMBER: US/09/826,660

; PRIOR FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 09/178,252

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: 60/065,215

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/076,445

; PRIOR FILING DATE: 1998-03-02

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 7

; LENGTH: 1815

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic B.t. toxin gene

US-09-826-660-7

Query Match 28.2% Score 25.4; DB 10; Length 1815;  
Best Local Similarity 68.6%; Pred. No. 12;  
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 3 TGACCTCTGTGATGATCATTCATTTGGGATCGGCAATGGCGACGACAT 53  
DB 430 TGATGCTGTGATGATCAAGCATCATCTGTGTAGCAAAAGCGATGCGACAT 380

RESULT 8

US-09-826-660-8/c

; Sequence 8, Application US/09826660

; Patent No. US20010026940A1

; GENERAL INFORMATION:

; APPLICANT: Cardineau, Guy A.

; APPLICANT: Steilman, Steven J.

; APPLICANT: Narva, Kenneth E.

; TITLE OF INVENTION: Plant-Optimized Genes Encoding Pesticidal Toxins

; FILE REFERENCE: NA-714XC2D1

; CURRENT APPLICATION NUMBER: US/09/826,660

; PRIOR FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 09/178,252

; PRIOR FILING DATE: 1998-10-23

; PRIOR APPLICATION NUMBER: 60/065,215

; PRIOR FILING DATE: 1997-11-12

; PRIOR APPLICATION NUMBER: 60/076,445

; PRIOR FILING DATE: 1998-03-02

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: Patentln Ver. 2.0

; SEQ ID NO 8

; LENGTH: 1815

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic B.t. toxin gene

US-09-826-660-8

Query Match 28.2% Score 25.4; DB 10; Length 1815;  
Best Local Similarity 68.6%; Pred. No. 12;  
Matches 35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 3 TGACCTCTGTGATGATCATTCATTTGGGATCGGCAATGGCGACGACAT 53  
DB 430 TGATGCTGTGATGATCAAGCATCATCTGTGTAGCAAAAGCGATGCGACAT 380

RESULT 9

US-09-969-373-1375/c

; Sequence 1375, Application US/09969373

; Patent No. US20020133852A1

; GENERAL INFORMATION:

; APPLICANT: Effertz, Roger J.

; APPLICANT: Haugue, Brian M.

; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping

; FILE REFERENCE: 38-10(52679)A

; CURRENT APPLICATION NUMBER: US/09/969,373

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 09/754,853

; PRIOR FILING DATE: 2001-01-05

; PRIOR APPLICATION NUMBER: US 09/760,427

; PRIOR FILING DATE: 2001-01-13

; PRIOR APPLICATION NUMBER: US 09/855,768

; PRIOR FILING DATE: 2001-05-15

; NUMBER OF SEQ ID NOS: 4593

; SEQ ID NO 1375

; LENGTH: 182

; TYPE: DNA

; ORGANISM: Glycine max

US-09-969-373-1375

Query Match 28.0% Score 25.2; DB 10; Length 182;  
Best Local Similarity 62.9%; Pred. No. 6.3;  
Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 13 GATGATCATCATTCATTTGGGATCGGCAATGGCGACGACATTAATCCACATTC 72  
DB 82 GAATGAATACATCATTCATTTGAATGTTAATGTTAAATAACCAATCCACGTC 23  
QY 73 AA 74  
DB 22 AA 21



```
Sequence 9565, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haseldeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9565
LENGTH: 2703
TYPE: DNA
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(2703)
US-09-815-242-9565

Query Match          30.4%; Score 27.4; DB 10; Length 2703;
Best Local Similarity 57.6%; Pred. No. 2.6;
Matches 49; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 2 GTGACCTGTGATGCAATCATCTGGGATCGCGATGGCAGCGACATPAATCCA 61
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2525 GGTATATCCGTGATGCGTGTATCTATGATGCGCACTCGAAGCTTGAACACTACA 2584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 AATCCACTTGAAACCACTGAAGAA 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2585 AAGACGACGTGAAGAAGATGACAAA 2609

RESULT 3
US-09-880-192-41
Sequence 41, Application US/09880192
Patent No. US20020077470A1
GENERAL INFORMATION:
APPLICANT: Walker, Michael G.
APPLICANT: Klingler, Todd M.
APPLICANT: Azimzal, Yalda
TITLE OF INVENTION: POLYNUCLEOTIDES ASSOCIATED WITH CARDIAC MUSCLE FUNCTION
FILE REFERENCE: PB-0009-1 CIP
CURRENT APPLICATION NUMBER: US/09/880,192
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 62
SOFTWARE: PERL Program
SEQ ID NO 41
LENGTH: 3441
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020077470A1 4151935CB1
US-09-880-192-41

Query Match          28.9%; Score 26; DB 10; Length 3441;
Best Local Similarity 76.2%; Pred. No. 9.4;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATAAATCCAAATCCACTTGACCACTGAAGAAAG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 775 GAACCTAAATCTGCTCCACCTGACGCAAGTGAAGAGAG 816

RESULT 4
US-10-114-893-126
Sequence 126, Application US/10114893
Publication No. US20020193567A1
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Collins-Racie, Lisa A.
APPLICANT: Evans, Cheryl
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Bowman, Michael R.
APPLICANT: Spaulding, Vikki
APPLICANT: Carlin-Duckett, McKeough
APPLICANT: Kelleher, Kerry S.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6000-10A
CURRENT APPLICATION NUMBER: US/10/114,893
CURRENT FILING DATE: 2002-04-02
EARLIER APPLICATION NUMBER: 09/413,232
EARLIER FILING DATE: 1999-10-06
NUMBER OF SEQ ID NOS: 321
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 126
LENGTH: 3444
TYPE: DNA
ORGANISM: Homo sapiens
US-10-114-893-126

Query Match          28.9%; Score 26; DB 9; Length 3444;
Best Local Similarity 76.2%; Pred. No. 9.4;
Matches 32; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 48 GCACATAAATCCAAATCCACTTGACCACTGAAGAAAG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 809 GAACCTAAATCTGCTCCACCTGACGCAAGTGAAGAGAG 850

RESULT 5
US-10-091-504-2256
Sequence 2256, Application US/10091504
Publication No. US20030059908A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC007C1
CURRENT APPLICATION NUMBER: US/10/091,504
CURRENT FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 2442
Prior Application removed - See File Wrapper or Palm
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2256
LENGTH: 4346
TYPE: DNA
ORGANISM: Homo sapiens
US-10-091-504-2256
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 15:55:39 ; Search time 10.3895 Seconds  
(without alignments)  
8702.120 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 90

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAGAGAGAAAGA 90

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 639749 seqs, 502280978 residues

Total number of hits satisfying chosen parameters: 1279498

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PTL\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCRTS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27.4	30.4	2000	US-09-938-842A-3043	Sequence 3043, Ap
2	27.4	30.4	2703	US-09-815-242-9565	Sequence 9565, Ap
3	26	28.9	3441	US-09-860-182-41	Sequence 41, Appl
4	26	28.9	3444	US-10-114-893-126	Sequence 126, Ap
5	26	28.9	4346	US-10-091-504-2256	Sequence 2256, Ap
6	26	28.9	4346	US-09-764-869-2256	Sequence 2256, Ap
7	25.4	28.2	1815	US-09-826-660-7	Sequence 7, Appl
8	25.4	28.2	1815	US-09-826-660-8	Sequence 8, Appl
9	25.2	28.0	182	US-09-969-373-1375	Sequence 1375, Ap
10	25.2	28.0	502	US-09-972-714-2	Sequence 2, Appl
11	25.2	28.0	1554	US-09-910-600-5	Sequence 5, Appl
12	25.2	28.0	1592	US-09-972-714-9	Sequence 9, Appl
13	25.2	28.0	6101	US-09-972-714-1	Sequence 1, Appl
14	24.8	27.6	653	US-09-764-860-1128	Sequence 1128, Ap
15	24.8	27.6	1529	US-09-925-302-212	Sequence 212, Ap
16	24.6	27.3	5470	US-09-764-864-1593	Sequence 1593, Ap
17	24.4	27.1	483	US-09-943-718-1	Sequence 1, Appl
18	24.4	27.1	912	US-09-792-793A-70	Sequence 70, Appl
19	24.4	27.1	1042	US-10-066-500-77	Sequence 77, Appl

C 20	24.4	27.1	1042	9	US-10-002-796-77	Sequence 77, Appl
C 21	24.4	27.1	1042	9	US-10-066-227-77	Sequence 77, Appl
C 22	24.4	27.1	1042	9	US-10-066-494-77	Sequence 77, Appl
C 23	24.4	27.1	1042	9	US-10-066-269-77	Sequence 77, Appl
C 24	24.4	27.1	1042	9	US-10-066-193-77	Sequence 77, Appl
C 25	24.4	27.1	1042	9	US-10-066-211-77	Sequence 77, Appl
C 26	24.2	26.9	1248	9	US-09-938-842A-266	Sequence 266, Ap
C 27	24.2	26.7	4629	10	US-09-150-811-7	Sequence 150, Ap
C 28	24	26.7	4999	9	US-10-007-968-14	Sequence 14, Appl
C 29	24	26.7	4999	12	US-09-740-211-14	Sequence 14, Appl
C 30	24	26.7	7944	12	US-10-095-718-1	Sequence 1, Appl
C 31	24	26.7	9009	9	US-09-957-641-1	Sequence 1, Appl
C 32	24	26.7	9009	9	US-10-187-319-1	Sequence 1, Appl
C 33	24	26.7	9029	9	US-10-132-828-1	Sequence 1, Appl
C 34	24	26.7	11933	9	US-10-007-968-13	Sequence 13, Appl
C 35	24	26.7	11933	10	US-09-740-211-13	Sequence 13, Appl
C 36	23.8	26.4	396	10	US-09-960-352-11737	Sequence 11737, A
C 37	23.8	26.4	462	9	US-09-925-299-335	Sequence 335, Ap
C 38	23.8	26.4	462	10	US-09-925-299-335	Sequence 335, Ap
C 39	23.8	26.4	818	9	US-10-202-193-116	Sequence 116, Appl
C 40	23.8	26.4	1037	9	US-09-974-879-53	Sequence 53, Appl
C 41	23.8	26.4	1815	10	US-09-826-660-3	Sequence 3, Appl
C 42	23.8	26.4	3444	10	US-09-826-660-1	Sequence 1, Appl
C 43	23.6	26.2	424	10	US-09-983-965-4039	Sequence 4039, Ap
C 44	23.6	26.2	858	9	US-09-738-626-3057	Sequence 3057, Ap
C 45	23.6	26.2	2061	9	US-09-938-842A-1375	Sequence 1375, Ap

## ALIGNMENTS

RESULT 1  
US-09-938-842A-3043  
Sequence 3043, Application US/0993842A  
Patent No. US20020160378A1  
GENERAL INFORMATION:  
APPLICANT: Harper, Jeff  
APPLICANT: Kleps, Joel  
APPLICANT: Wang, Xun  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
TITLE OF INVENTION: SAME AND METHODS OF USE  
FILE REFERENCE: SRRIP1300-3  
CURRENT APPLICATION NUMBER: US/09/938, 842A  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: US 60/227, 866  
PRIOR FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: US 60/264, 647  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/300, 111  
PRIOR FILING DATE: 2001-06-22  
NUMBER OF SEQ ID NOS: 5379  
SEQ ID NO 3043  
LENGTH: 2000  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-938-842A-3043  
Query Match  
Best Local Similarity 30.4%; Score 27.4; DB 9; Length 2000;  
Matches 49; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
QY 2 GTGACCTCTGTGATCGAATCATTCATTTGGGATCGCAATGGCGAGCGACATTAATCA 61  
DB 337 GAGCCATCGAGATCTGTGATCATTCATTTAGCCGCGACATGGCGACCATTCCTTAACA 396  
QY 62 AATCCACTGGAACGACGAGAA 86  
DB 397 CATCTCTTCAAGCAGATGACAAA 421  
RESULT 2  
US-09-815-242-9565

Query Match 100.0%; Score 90; DB 17; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 60  
|||||  
DB 923 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 864  
OY 61 AATCCCACTTGAACCACTGAAGAAAGA 90  
|||||  
DB 863 AATCCCACTTGAACCACTGAAGAAAGA 834

RESULT 14  
US-09-898-888-8140/c  
; Sequence 8140, Application US/09898888  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
; FILE REFERENCE: 20411-748CON1  
; CURRENT APPLICATION NUMBER: US/09/898,888  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 09/340,623  
; NUMBER OF SEQ ID NOS: 45207  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8140  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-898-888-8140

Query Match 100.0%; Score 90; DB 33; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 60  
|||||  
DB 923 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 864  
OY 61 AATCCCACTTGAACCACTGAAGAAAGA 90  
|||||  
DB 863 AATCCCACTTGAACCACTGAAGAAAGA 834

RESULT 15  
US-09-898-888A-8140/c  
; Sequence 8140, Application US/09898888A  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
; FILE REFERENCE: 20411-748CON1  
; CURRENT APPLICATION NUMBER: US/09/898,888A  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/340,623  
; PRIOR FILING DATE: 1999-06-28  
; PRIOR APPLICATION NUMBER: US 09/205,070  
; NUMBER OF SEQ ID NOS: 45207  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8140  
; LENGTH: 1041  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-898-888A-8140

Query Match 100.0%; Score 90; DB 33; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 60  
|||||  
DB 923 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGGCGACGCCACATAAATCC 864  
OY 61 AATCCCACTTGAACCACTGAAGAAAGA 90  
|||||  
DB 863 AATCCCACTTGAACCACTGAAGAAAGA 834

Search completed: April 21, 2003, 17:14:11  
Job time : 267.819 secs

;; CURRENT FILING DATE: 1999-07-22  
;; EARLIER APPLICATION NUMBER: US 09/205,155  
;; EARLIER FILING DATE: 1998-12-03  
;; NUMBER OF SEQ ID NOS: 13203  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3465  
;; LENGTH: 1039  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-359-922-3465

Query Match 100.0%; Score 90; DB 17; Length 1039;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60  
|||  
DB 921 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 862  
|||  
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||  
DB 861 AATCCCACTTGAACCACTGAAGAAAAGA 832

RESULT 10  
US-09-359-922-3465/c  
;; Sequence 3465, Application US/09359922A  
;; GENERAL INFORMATION:  
;; APPLICANT: Leshkowitz, Dena  
;; APPLICANT: Liu, Jin  
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
;; FILE REFERENCE: 20411-752CON1  
;; CURRENT APPLICATION NUMBER: US/09/359,922A  
;; CURRENT FILING DATE: 1999-07-22  
;; EARLIER APPLICATION NUMBER: US 09/205,155  
;; EARLIER FILING DATE: 1998-12-03  
;; EARLIER APPLICATION NUMBER: US 09/034,341  
;; EARLIER FILING DATE: 1998-02-13  
;; NUMBER OF SEQ ID NOS: 13203  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3465  
;; LENGTH: 1039  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-359-922-3465

Query Match 100.0%; Score 90; DB 17; Length 1039;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60  
|||  
DB 921 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 862  
|||  
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||  
DB 861 AATCCCACTTGAACCACTGAAGAAAAGA 832

RESULT 11  
US-09-919-002-3465/c  
;; Sequence 3465, Application US/09919002  
;; GENERAL INFORMATION:  
;; APPLICANT: Leshkowitz, Dena  
;; APPLICANT: Liu, Jin  
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
;; FILE REFERENCE: 20411-752CON1  
;; CURRENT APPLICATION NUMBER: US/09/919,002  
;; CURRENT FILING DATE: 2001-07-30  
;; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: 09/359,922  
;; PRIOR FILING DATE: FILING DATE: 1999-07-22

;; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US 09/034,341  
;; PRIOR FILING DATE: FILING DATE: 1998-02-13  
;; NUMBER OF SEQ ID NOS: 13203  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 3465  
;; LENGTH: 1039  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-919-002-3465

Query Match 100.0%; Score 90; DB 34; Length 1039;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60  
|||  
DB 921 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 862  
|||  
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||  
DB 861 AATCCCACTTGAACCACTGAAGAAAAGA 832

RESULT 12  
US-09-205-070-8140/c  
;; Sequence 8140, Application US/09205070  
;; GENERAL INFORMATION:  
;; APPLICANT: Hysq, Inc.  
;; APPLICANT: Hysq, Inc.  
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
;; FILE REFERENCE: 20411-748  
;; CURRENT APPLICATION NUMBER: US/09/205,070  
;; CURRENT FILING DATE: 1998-12-03  
;; NUMBER OF SEQ ID NOS: 45207  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8140  
;; LENGTH: 1041  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-205-070-8140

Query Match 100.0%; Score 90; DB 16; Length 1041;  
Best Local Similarity 100.0%; Pred. No. 1.5e-21;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60  
|||  
DB 923 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 864  
|||  
OY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||  
DB 863 AATCCCACTTGAACCACTGAAGAAAAGA 834

RESULT 13  
US-09-340-623-8140/c  
;; Sequence 8140, Application US/09340623  
;; GENERAL INFORMATION:  
;; APPLICANT: Hysq, Inc.  
;; APPLICANT: Hysq, Inc.  
;; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA  
;; FILE REFERENCE: 20411-748CON1  
;; CURRENT APPLICATION NUMBER: US/09/340,623  
;; CURRENT FILING DATE: 1999-06-28  
;; EARLIER APPLICATION NUMBER: US 09/205,070  
;; EARLIER FILING DATE: 1998-12-03  
;; NUMBER OF SEQ ID NOS: 45207  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8140  
;; LENGTH: 1041  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-340-623-8140

```
OY      1  TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGCAGCCACATAAATCC 60
          |||||||
DB      278 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGCAGCCACATAAATCC 337
OY      61  AAATCCCACTTGACCACTGAAGAAAAGA 90
          |||||||
DB      338 AAATCCCACTTGACCACTGAAGAAAAGA 367
```

## RESULT 7

```
US-09-515-128-21383
; Sequence 21383, Application US/09515128
; GENERAL INFORMATION:
```

```
; APPLICANT: Atterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroza, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogira, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralsi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/515,128
; CURRENT FILING DATE: 2000-02-28
; EARLIER APPLICATION NUMBER: 09/034,341
; EARLIER FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21383
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(442)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-09-515-128-21383
```

```
Query Match          100.0%; Score 90; DB 19; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGCAGCCACATAAATCC 60
          |||||||
DB      149 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGCAGCCACATAAATCC 208
OY      61  AAATCCCACTTGACCACTGAAGAAAAGA 90
          |||||||
DB      209 AAATCCCACTTGACCACTGAAGAAAAGA 238
```

## RESULT 8

```
US-09-721-544-21383
```

```
; Sequence 21383, Application US/09721544
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Atterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Hunter, Kelly
; APPLICANT: Jessen, Aaron
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroza, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogira, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Ralsi, Fariba
; APPLICANT: Smith, Benjamin
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Yang, Fel
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; TITLE OF INVENTION: From a cDNA Library of Fetal Liver-Spleen
; FILE REFERENCE: 728CIP
; CURRENT APPLICATION NUMBER: US/09/721,544
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 09/515,128
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: 09/034,341
; PRIOR FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 24489
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21383
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(442)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-09-721-544-21383
```

```
Query Match          100.0%; Score 90; DB 29; Length 442;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGCAGCCACATAAATCC 60
          |||||||
DB      149 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGCAGCCACATAAATCC 208
OY      61  AAATCCCACTTGACCACTGAAGAAAAGA 90
          |||||||
DB      209 AAATCCCACTTGACCACTGAAGAAAAGA 238
```

## RESULT 9

```
US-09-359-922-3465/C
; Sequence 3465, Application US/09359922
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Hysed, Inc.
; TITLE OF INVENTION: NOVEL COUNTS OBTAINED FROM VARIOUS cDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922
```

```
APPLICANT: Naughton, Rebecca E.
APPLICANT: Klingler, Tod M.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM
TITLE OF INVENTION: HUMAN FETAL LUNG
NUMBER OF SEQUENCES: 2102
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/60/044,082
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PD-0357P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 855-0555
TELEFAX: (415) 845-4166
INFORMATION FOR SEQ ID NO: 268:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: 2888829H1
US-60-044-082-268

Query Match
Best Local Similarity 100.0%; Score 90; DB 48; Length 263;
Matches 90; Conservative 0; Mismatches 1e-21; Indels 0; Gaps 0;

QY 1 TGTACCTGTGTGATCATCATTTGGGATGCGGATGGGACCGCATATAAATCC 60
Db 34 TGTACCTGTGTGATCATCATTTGGGATGCGGATGGGACCGCATATAAATCC 93

QY 61 AATCCCACTTGAACCACTGAAGAAAGA 90
Db 94 AATCCCACTTGAACCACTGAAGAAAGA 123

RESULT 4
US-09-532-315-9519
Sequence 9519, Application US/09532315
GENERAL INFORMATION:
APPLICANT: Sellhammer, Jeffrey J.
APPLICANT: Deleageane, Angelo M.
APPLICANT: Stuart, Susan G.
APPLICANT: Stuve, Laura L.
APPLICANT: Mullahy, Sara J.
APPLICANT: Naughton, Rebecca E.
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES
FILE REFERENCE: PD-1002 CIP
CURRENT APPLICATION NUMBER: US/09/532,315
CURRENT FILING DATE: 2000-03-24
Pilot application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 42212
SOFTWARE: PERL Program
SEQ ID NO 9519
LENGTH: 289
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No: hu00815059
US-09-532-315-9519

Query Match
Best Local Similarity 100.0%; Score 90; DB 20; Length 289;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTACCTGTGTGATCATCATTTGGGATGCGGATGGGACCGCATATAAATCC 60
Db 84 TGTACCTGTGTGATCATCATTTGGGATGCGGATGGGACCGCATATAAATCC 143

QY 61 AATCCCACTTGAACCACTGAAGAAAGA 90
Db 144 AATCCCACTTGAACCACTGAAGAAAGA 173

RESULT 5
US-09-489-036-7332
Sequence 7332, Application US/09489036
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 783
CURRENT APPLICATION NUMBER: US/09/489,036
CURRENT FILING DATE: 2000-01-19
NUMBER OF SEQ ID NOS: 35324
SOFTWARE: Hy-patent.pl Version 3.1
SEQ ID NO 7332
LENGTH: 372
TYPE: DNA
ORGANISM: Homo sapiens
US-09-489-036-7332

Query Match
Best Local Similarity 100.0%; Score 90; DB 18; Length 372;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTACCTGTGTGATCATCATTTGGGATGCGGATGGGACCGCATATAAATCC 60
Db 278 TGTACCTGTGTGATCATCATTTGGGATGCGGATGGGACCGCATATAAATCC 337

QY 61 AATCCCACTTGAACCACTGAAGAAAGA 90
Db 338 AATCCCACTTGAACCACTGAAGAAAGA 367

RESULT 6
US-09-943-143-7332
Sequence 7332, Application US/09943143
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
FILE REFERENCE: 783
CURRENT APPLICATION NUMBER: US/09/943,143
CURRENT FILING DATE: 2001-08-30
Pilot application data removed - refer to PALM or file wrapper
PRIORITY FILING DATE: 2000-01-19
NUMBER OF SEQ ID NOS: 35324
SOFTWARE: Hy-patent.pl Version 3.1
SEQ ID NO 7332
LENGTH: 372
TYPE: DNA
ORGANISM: Homo sapiens
US-09-943-143-7332

Query Match
Best Local Similarity 100.0%; Score 90; DB 35; Length 372;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

22 90 100.0 1858 18 US-09-471-275-2814
23 90 100.0 2041 19 US-09-513-151-3
24 90 100.0 2101 61 US-60-172-360-26034
25 90 100.0 2108 40 US-10-133-013-165
26 90 100.0 2109 76 US-60-324-185-33690
27 90 100.0 2129 1 PCT-US02-07826-150
28 90 100.0 2129 1 PCT-US02-07826-150
29 90 100.0 2129 39 US-10-097-340-152
30 90 100.0 2129 39 US-10-097-340-152
31 88.4 98.2 1069 25 US-09-652-124-7628
32 75.4 83.8 287 12 US-08-810-326-938
33 75.4 83.8 287 20 US-09-533-315-9504
34 75.4 83.8 287 45 US-60-012-699-938
35 74 82.2 2103 80 US-60-360-207-2119
36 72.4 80.4 301 15 US-09-151-199-3234
37 72.4 80.4 301 18 US-09-482-965-1320
38 72.4 80.4 301 20 US-09-539-806-4638
39 56.4 62.7 523 23 US-09-617-081-981
40 52.8 58.7 434 11 US-08-798-074-1079
41 52.8 58.7 434 11 US-08-798-074B-1079
42 49.4 54.9 51 76 US-60-324-185-33691
43 47.2 52.4 221 35 US-09-880-093-760
44 47.2 52.4 248 13 US-09-107-425-1617
45 47.2 52.4 248 21 US-09-540-764-25285

```

```

Sequence 2814, Ap
Sequence 3, Appli
Sequence 26034, A
Sequence 165, App
Sequence 33690, A
Sequence 150, App
Sequence 152, App
Sequence 152, App
Sequence 152, App
Sequence 7628, Ap
Sequence 938, App
Sequence 938, App
Sequence 938, App
Sequence 2119, Ap
Sequence 3234, Ap
Sequence 1320, Ap
Sequence 4638, A
Sequence 981, App
Sequence 1079, Ap
Sequence 1079, Ap
Sequence 33691, A
Sequence 760, App
Sequence 1617, Ap
Sequence 25285, A

```

## ALIGNMENTS

## RESULT 1

US-09-056-942-268

Sequence 268, Application US/09056942

GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.

APPLICANT: Stuve, Laura L.

APPLICANT: Stuart, Susan G.

APPLICANT: Ito, Laura Y.

APPLICANT: Akerblom, Ingrid E.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Naughton, Rebecca E.

APPLICANT: Klingler, Tod M.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM

TITLE OF INVENTION: HUMAN FETAL LUNG

NUMBER OF SEQUENCES: 2102

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/056,942

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/044,082

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: CERRONE, MICHAEL C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PD-0357P

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 855-0555

TELEFAX: (415) 845-4166

INFORMATION FOR SEQ ID NO: 268:

SEQUENCE CHARACTERISTICS:

LENGTH: 263 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: 2888829H1
US-09-056-942-268

```

```

Query Match 100.0%; Score 90; DB 14; Length 263;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TGTGACCTCTGTGATGATCATCATTTGGGATCGGAATGGCGACGACATTAATTC 60
DB 34 TGTGACCTCTGTGATGATCATCATTTGGGATCGGAATGGCGACGACATTAATTC 93
DB 94 AAATCCACTTGAACCACTGAAGAAAGA 123

```

## RESULT 2

US-09-532-315-9510

Sequence 9510, Application US/09532315

GENERAL INFORMATION:

APPLICANT: Selhammer, Jeffrey J.

APPLICANT: Delegeane, Angelo M.

APPLICANT: Stuart, Susan G.

APPLICANT: Stuve, Laura L.

APPLICANT: Mullahy, Sara J.

APPLICANT: Naughton, Rebecca E.

TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES

FILE REFERENCE: PD-1002 CIP

CURRENT APPLICATION NUMBER: US/09/532,315

CURRENT FILING DATE: 2000-03-24

Prior application data removed - refer to PALM or file wrapper

NUMBER OF SEQ ID NOS: 42212

SOFTWARE: PERL Program

SEQ ID NO 9510

LENGTH: 263

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

OTHER INFORMATION: Incyte ID No: hu00416615

NAME/KEY: unsure

LOCATION: 228, 241

OTHER INFORMATION: a, t, c, g, or other

US-09-532-315-9510

```

Query Match 100.0%; Score 90; DB 20; Length 263;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TGTGACCTCTGTGATGATCATCATTTGGGATCGGAATGGCGACGACATTAATTC 60
DB 34 TGTGACCTCTGTGATGATCATCATTTGGGATCGGAATGGCGACGACATTAATTC 93
QY 61 AAATCCACTTGAACCACTGAAGAAAGA 90
DB 94 AAATCCACTTGAACCACTGAAGAAAGA 123

```

## RESULT 3

US-60-044-082-268

Sequence 268, Application US/60044082

GENERAL INFORMATION:

APPLICANT: Gooding, Douglas H.

APPLICANT: Stuve, Laura L.

APPLICANT: Stuart, Susan G.

APPLICANT: Ito, Laura Y.

APPLICANT: Akerblom, Ingrid E.

APPLICANT: Delegeane, Angelo M.

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:26:48 ; Search time 265.819 Seconds

(without alignments)  
8512.654 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 90  
Sequence: 1 TGTGACCTGTGTGATCGAAT.....TGAACCACTGAAGAAAAGA 90

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 08  
Listing first 45 summaries

Database :

Pending Patents, NA, Main: \*

```

1: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

```

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	263	14	US-09-056-942-268
2	90	100.0	263	20	US-09-532-315-9510
3	90	100.0	263	48	US-60-044-082-268
4	90	100.0	289	20	US-09-532-315-9519
5	90	100.0	372	18	US-09-489-036-7332
6	90	100.0	372	35	US-09-943-143-7332
7	90	100.0	442	19	US-09-515-188-21383
8	90	100.0	442	29	US-09-721-544-21383
9	90	100.0	1039	17	US-09-359-922-3465
10	90	100.0	1039	17	US-09-359-922-3465
11	90	100.0	1039	34	US-09-919-002-8146
12	90	100.0	1041	16	US-09-205-070-8140
13	90	100.0	1041	17	US-09-340-623-8140
14	90	100.0	1041	33	US-09-898-888-8140
15	90	100.0	1041	33	US-09-898-888-8140
16	90	100.0	1183	29	US-09-757-028-874
17	90	100.0	1183	42	US-10-232-911-874
18	90	100.0	1223	28	US-09-705-256A-6006
19	90	100.0	1223	60	US-60-164-285-6006
20	90	100.0	1749	1	PCT-US02-18947-1692
21	90	100.0	1749	41	US-10-172-118-1692

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES



Db 1211 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGCATGGCAGCCACATATAATCC 1270  
OY 61 AATCCCACTTGAACCACTGAGAGAAAAGA 90  
Db 1271 AATCCCACTTGAACCACTGAGAGAAAAGA 1300

## RESULT 14

US-10-380-731-119  
Sequence 119, Application US/10380731  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 21272-114  
CURRENT APPLICATION NUMBER: US/10/380,731  
CURRENT FILING DATE: 2003-03-14  
PRIOR APPLICATION NUMBER: 09/659,671  
PRIOR FILING DATE: 2000-09-11  
NUMBER OF SEQ ID NOS: 888  
SOFTWARE: Custom  
SEQ ID NO 119  
LENGTH: 2130  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (191)..(1426)  
US-10-380-731-119

Query Match 100.0%; Score 90; DB 8; Length 2130;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGCATGGCAGCCACATATAATCC 60  
Db 1211 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGCATGGCAGCCACATATAATCC 1270  
OY 61 AATCCCACTTGAACCACTGAGAGAAAAGA 90  
Db 1271 AATCCCACTTGAACCACTGAGAGAAAAGA 1300

## RESULT 15

US-09-532-315B-9504  
Sequence 9504, Application US/09532315B  
GENERAL INFORMATION:  
APPLICANT: Sellhamey, Jeffrey J.  
APPLICANT: Deleage, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullishy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES  
FILE REFERENCE: PD-1002 CIP  
CURRENT APPLICATION NUMBER: US/09/532,315B  
CURRENT FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 07/916,491  
PRIOR FILING DATE: 1992-07-17  
PRIOR APPLICATION NUMBER: 07/977,780  
PRIOR FILING DATE: 1992-11-19  
PRIOR APPLICATION NUMBER: 08/100,523  
PRIOR FILING DATE: 1993-08-03  
PRIOR APPLICATION NUMBER: 09/008,119  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 08/196,364  
PRIOR FILING DATE: 1994-02-14  
PRIOR APPLICATION NUMBER: 08/282,991  
PRIOR FILING DATE: 1994-07-28  
PRIOR APPLICATION NUMBER: 08/438,571  
PRIOR FILING DATE: 1995-05-10  
PRIOR APPLICATION NUMBER: 08/179,873  
PRIOR FILING DATE: 1994-01-11

PRIOR APPLICATION NUMBER: 08/504,732  
PRIOR FILING DATE: 1995-07-20  
PRIOR APPLICATION NUMBER: 08/137,951  
PRIOR FILING DATE: 1993-10-14  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 42212  
SOFTWARE: PERL Program  
SEQ ID NO 9504  
LENGTH: 287  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc-feature  
OTHER INFORMATION: Incyte ID No: hu00221292  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 107, 113, 120-121, 174, 259, 279, 282  
OTHER INFORMATION: a, t, c, g, or other  
US-09-532-315B-9504

Query Match 83.8%; Score 75.4; DB 6; Length 287;  
Best Local Similarity 96.7%; Pred. No. 2.3e-17;  
Matches 87; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGC-GAATGGCAGCCACATATAATC 59  
Db 178 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCAGCCACATATAATC 237  
OY 60 CAATCCCACTTGAACCACTGAGAGAAAAG 89  
Db 238 CAATCCCACTTGAACCACTGAGAGAAAAG 267

Search completed: April 21, 2003, 18:41:28  
Job time: 52.7649 secs

US-10-342-887-1692

Query Match 100.0%; Score 90; DB 9; Length 1749;  
Best Local Similarity 100.0%; Pred. No. 1.7e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 60  
DB 826 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 885  
QY 61 AATCCCACTTGACCAACTGAAGAAAGA 90  
DB 886 AATCCCACTTGACCAACTGAAGAAAGA 915

RESULT 10  
US-10-170-235-40554  
Sequence 40554, Application US/10170235  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig  
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN  
TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF  
FILE REFERENCE: C1001380  
CURRENT APPLICATION NUMBER: US/10/170,235  
CURRENT FILING DATE: 2003-03-17  
NUMBER OF SEQ ID NOS: 42514  
SEQ ID NO 40554  
LENGTH: 2119  
TYPE: DNA  
ORGANISM: HUMAN  
US-10-170-235-40554

Query Match 100.0%; Score 90; DB 8; Length 2119;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 60  
DB 1213 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 1272  
QY 61 AATCCCACTTGACCAACTGAAGAAAGA 90  
DB 1273 AATCCCACTTGACCAACTGAAGAAAGA 1302

RESULT 11  
US-60-453-135-275  
Sequence 275, Application US/60453135  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: IAKOUBOVA, Olga  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001456  
CURRENT APPLICATION NUMBER: US/60/453,135  
CURRENT FILING DATE: 2003-03-10  
NUMBER OF SEQ ID NOS: 82762  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 275  
LENGTH: 2119  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-60-453-135-275

Query Match 100.0%; Score 90; DB 11; Length 2119;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 60  
DB 1213 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 1272  
QY 61 AATCCCACTTGACCAACTGAAGAAAGA 90

DB 1273 AATCCCACTTGACCAACTGAAGAAAGA 1302

RESULT 12  
US-60-453-050-275  
Sequence 275, Application US/60453050  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: LIKE, May  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: C1001457  
CURRENT APPLICATION NUMBER: US/60/453,050  
CURRENT FILING DATE: 2003-03-10  
NUMBER OF SEQ ID NOS: 82762  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 275  
LENGTH: 2119  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-60-453-050-275

Query Match 100.0%; Score 90; DB 11; Length 2119;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 60  
DB 1213 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 1272  
QY 61 AATCCCACTTGACCAACTGAAGAAAGA 90  
DB 1273 AATCCCACTTGACCAACTGAAGAAAGA 1302

RESULT 13  
US-09-659-671A-134  
Sequence 134, Application US/09659671A  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Zhou, Ping  
APPLICANT: Asundi, Vinod  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhang, Jie  
APPLICANT: Zhao, Qing A.  
APPLICANT: Ren, Feiyang J.  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wehtman, Tom  
APPLICANT: Dimanac, Radoje T.  
TITLE OF INVENTION: Novel Nucleic Acids and  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 794  
CURRENT APPLICATION NUMBER: US/09/659,671A  
CURRENT FILING DATE: 2000-09-11  
NUMBER OF SEQ ID NOS: 475  
SOFTWARE: PL-FL-genes Version 2.0  
SEQ ID NO 134  
LENGTH: 2130  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-659-671A-134

Query Match 100.0%; Score 90; DB 6; Length 2130;  
Best Local Similarity 100.0%; Pred. No. 1.8e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATTCATTTGGGATCGCGAATGGGCGACGACATAAATCC 60

```

; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 276
; LENGTH: 1271
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-276

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 1271;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60
Db 365 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 424

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
Db 425 AATCCCACTTGAACCACTGAAGAAAAGA 454

RESULT 6
US-10-170-235-41123
; Sequence 41123, Application US/10170235
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN
; FILE REFERENCE: CL001380
; CURRENT APPLICATION NUMBER: US/10/170,235
; CURRENT FILING DATE: 2003-03-17
; NUMBER OF SEQ ID NOS: 42514
; SEQ ID NO 41123
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: HUMAN
US-10-170-235-41123

Query Match
Best Local Similarity 100.0%; Score 90; DB 8; Length 1719;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60
Db 813 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 872

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
Db 873 AATCCCACTTGAACCACTGAAGAAAAGA 902

RESULT 7
US-60-453-135-274
; Sequence 274, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-274

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 1719;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60
Db 813 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 872

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
Db 873 AATCCCACTTGAACCACTGAAGAAAAGA 902

RESULT 8
US-60-453-050-274
; Sequence 274, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-050-274

Query Match
Best Local Similarity 100.0%; Score 90; DB 11; Length 1719;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 60
Db 813 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGGCGAGCGCACATAAAATCC 872

QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90
Db 873 AATCCCACTTGAACCACTGAAGAAAAGA 902

RESULT 9
US-10-342-887-1692
; Sequence 1692, Application US/10342887
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1692
; LENGTH: 1749
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1692
```

NAME/KEY: unsure  
LOCATION: 228, 241  
OTHER INFORMATION: a, t, c, g, or other  
US-09-532-315B-9510

Query Match 100.0%; Score 90; DB 6; Length 263;  
Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 60  
|||||  
DB 34 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 93  
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||||  
DB 94 AATCCCACTTGAACCACTGAAGAAAAGA 123

#### RESULT 2

US-09-532-315B-9519

Sequence 9519, Application US/09532315B  
GENERAL INFORMATION:  
APPLICANT: Sellhammer, Jeffrey J.  
APPLICANT: Deleese, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES  
FILE REFERENCE: PD-1002 CIP  
CURRENT APPLICATION NUMBER: US/09/532.315B  
PRIORITY FILING DATE: 2000-03-24  
PRIORITY FILING DATE: 1992-07-17  
PRIORITY FILING DATE: 1992-07-17  
PRIORITY FILING DATE: 1992-11-19  
PRIORITY FILING DATE: 1992-11-19  
PRIORITY FILING DATE: 1993-08-03  
PRIORITY FILING DATE: 1993-08-03  
PRIORITY FILING DATE: 1998-01-16  
PRIORITY FILING DATE: 1998-01-16  
PRIORITY FILING DATE: 1994-02-14  
PRIORITY FILING DATE: 1994-02-14  
PRIORITY FILING DATE: 1994-07-28  
PRIORITY FILING DATE: 1994-07-28  
PRIORITY FILING DATE: 1995-05-10  
PRIORITY FILING DATE: 1995-05-10  
PRIORITY FILING DATE: 1994-01-11  
PRIORITY FILING DATE: 1994-01-11  
PRIORITY FILING DATE: 1995-07-20  
PRIORITY FILING DATE: 1995-07-20  
PRIORITY FILING DATE: 1993-10-14  
PRIORITY FILING DATE: 1993-10-14  
Remainder of prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 42212  
SOFTWARE: PERL Program  
SEQ ID NO 9519  
LENGTH: 289  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: hu00815059  
US-09-532-315B-9519

Query Match 100.0%; Score 90; DB 6; Length 289;  
Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 60  
|||||  
DB 84 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 143  
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90

|||||  
DB 144 AATCCCACTTGAACCACTGAAGAAAAGA 173

#### RESULT 3

US-10-170-235-41124

Sequence 41124, Application US/10170235

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig

TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HU

TITLE OF INVENTION: TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF

FILE REFERENCE: C1001380

CURRENT APPLICATION NUMBER: US/10/170,235

CURRENT FILING DATE: 2003-03-17

NUMBER OF SEQ ID NOS: 42514

SEQ ID NO 41124

LENGTH: 1271

TYPE: DNA

ORGANISM: HUMAN

US-10-170-235-41124

Query Match 100.0%; Score 90; DB 8; Length 1271;  
Best Local Similarity 100.0%; Pred. No. 1.6e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 60  
|||||  
DB 365 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 424  
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||||  
DB 425 AATCCCACTTGAACCACTGAAGAAAAGA 454

#### RESULT 4

US-60-453-135-276

Sequence 276, Application US/60453135

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: IAKOBBOVA, Olga

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001456

CURRENT APPLICATION NUMBER: US/60/453,135

CURRENT FILING DATE: 2003-03-10

NUMBER OF SEQ ID NOS: 82762

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 276

LENGTH: 1271

TYPE: DNA

ORGANISM: Homo sapiens

US-60-453-135-276

Query Match 100.0%; Score 90; DB 11; Length 1271;  
Best Local Similarity 100.0%; Pred. No. 1.6e-22;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 60  
|||||  
DB 365 TGTGACCTCTGTGATGCAATCATCATTTGGGGATGCGAATGGCGAGCGCACATAAAATCC 424  
QY 61 AATCCCACTTGAACCACTGAAGAAAAGA 90  
|||||  
DB 425 AATCCCACTTGAACCACTGAAGAAAAGA 454

#### RESULT 5

US-60-453-050-276

Sequence 276, Application US/60453050

GENERAL INFORMATION:

APPLICANT: CARGILL, Michele

APPLICANT: LUKE, May

TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 21, 2003, 13:37:18 ; Search time 50.7649 Seconds  
(without alignments)  
7947.624 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 90  
Sequence: 1 TGTGACCTGTGATCGAAT.....TGACCACTGAGAAAGA 90

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 5930009 seqs, 2241446263 residues

Total number of hits satisfying chosen parameters: 11860018

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Pending Patents, NA, New: \*  
1: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq: \*  
2: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq: \*  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq: \*  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq: \*  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq: \*  
6: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq: \*  
7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq: \*  
8: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq: \*  
9: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq: \*  
10: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq: \*  
11: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	263	6	US-09-532-315B-9510 Sequence 9510, Ap
2	90	100.0	289	6	US-09-532-315B-9519 Sequence 9519, Ap
3	90	100.0	1271	8	US-10-170-235-41124 Sequence 41124, A
4	90	100.0	1271	11	US-60-453-135-276 Sequence 276, App
5	90	100.0	1271	11	US-60-453-050-276 Sequence 276, App
6	90	100.0	1719	8	US-10-170-235-41123 Sequence 41123, A
7	90	100.0	1719	11	US-60-453-135-274 Sequence 274, App
8	90	100.0	1719	11	US-60-453-050-274 Sequence 274, App
9	90	100.0	1749	9	US-10-342-887-1692 Sequence 1692, Ap
10	90	100.0	2119	8	US-10-170-235-40554 Sequence 40554, A
11	90	100.0	2119	11	US-60-453-135-275 Sequence 275, App
12	90	100.0	2119	11	US-60-453-050-275 Sequence 275, App
13	90	100.0	2130	6	US-09-659-671A-134 Sequence 134, App
14	90	100.0	2130	6	US-10-380-731-119 Sequence 119, App
15	90	100.0	287	6	US-09-532-315B-9504 Sequence 9504, Ap
16	75.4	82.2	2103	7	US-10-144-771-2119 Sequence 2119, Ap
17	72.4	80.4	301	7	US-09-539-806B-46638 Sequence 46638, A
18	52.8	58.7	434	6	US-09-912-293-54990 Sequence 54990, A
19	47.2	52.4	248	9	US-10-349-781-405285 Sequence 25285, A
20	47	52.2	246	9	US-10-349-781-405285 Sequence 405285, A
21	29.8	33.1	241	6	US-09-532-315B-9512 Sequence 9512, Ap
22	28.8	32.0	1037984	7	US-09-947-911-259 Sequence 259, App

23	28	31.1	358	11	US-60-141-233-66580 Sequence 66580, A
24	27.6	30.7	259	7	US-09-837-604A-36790 Sequence 36790, A
25	27.6	30.7	441	7	US-09-837-604A-80102 Sequence 80102, A
26	27.6	30.7	458	7	US-09-837-604A-18601 Sequence 18601, A
27	27.6	30.7	462	7	US-09-837-604A-52531 Sequence 52531, A
28	27.6	30.7	519	7	US-09-837-604A-35887 Sequence 35887, A
29	27.6	30.7	593	7	US-09-837-604A-54885 Sequence 54885, A
30	27.6	30.7	605	7	US-09-837-604A-53708 Sequence 53708, A
31	27.6	30.7	22345	6	US-09-949-002-681 Sequence 681, App
32	27.6	30.7	22345	6	US-09-949-002-746 Sequence 746, App
33	27.6	30.7	22347	6	US-09-949-002-745 Sequence 745, App
34	26.8	29.8	599	8	US-10-115-678-136 Sequence 136, App
35	26.8	29.8	599	8	US-10-112-944-136 Sequence 136, App
36	26.6	29.6	601	7	US-09-949-016-162940 Sequence 162940, A
37	26.6	29.6	42118	7	US-09-949-016-16297 Sequence 16297, A
38	26.4	29.3	3474	9	US-10-144-771-5708 Sequence 5708, App
39	26.4	29.3	3687	9	US-10-144-771-21447 Sequence 21447, A
40	26.2	29.1	396	11	US-60-141-233-45678 Sequence 45678, A
41	26.2	29.1	807	6	US-09-134-000C-1231 Sequence 1231, Ap
42	26.2	29.1	807	6	US-09-134-000C-1231 Sequence 1231, Ap
43	26.2	29.1	24837	7	US-09-949-016-16177 Sequence 16177, A
44	26	28.9	200	6	US-09-532-315B-9517 Sequence 9517, Ap
45	26	28.9	201	11	US-60-453-135-65757 Sequence 65757, A

#### ALIGNMENTS

RESULT 1  
US-09-532-315B-9510  
Sequence 9510, Application US/09532315B  
GENERAL INFORMATION:  
APPLICANT: Sellhammer, Jeffrey J.  
APPLICANT: Deleage, Angelo M.  
APPLICANT: Stuart, Susan G.  
APPLICANT: Stuve, Laura L.  
APPLICANT: Mullahy, Sara J.  
APPLICANT: Naughton, Rebecca E.  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING TRANSFERASES  
FILE REFERENCE: PD-1002 CIP  
CURRENT APPLICATION NUMBER: US/09/532, 315B  
CURRENT FILING DATE: 2000-03-24  
PRIOR APPLICATION NUMBER: 07/916,491  
PRIOR FILING DATE: 1992-07-17  
PRIOR APPLICATION NUMBER: 07/977,780  
PRIOR FILING DATE: 1992-11-19  
PRIOR APPLICATION NUMBER: 08/100,523  
PRIOR FILING DATE: 1993-08-03  
PRIOR APPLICATION NUMBER: 09/008,119  
PRIOR FILING DATE: 1998-01-16  
PRIOR APPLICATION NUMBER: 08/196,364  
PRIOR FILING DATE: 1994-02-14  
PRIOR APPLICATION NUMBER: 08/282,991  
PRIOR FILING DATE: 1994-07-28  
PRIOR APPLICATION NUMBER: 08/438,571  
PRIOR FILING DATE: 1995-05-10  
PRIOR APPLICATION NUMBER: 08/179,873  
PRIOR FILING DATE: 1994-01-11  
PRIOR APPLICATION NUMBER: 08/504,732  
PRIOR FILING DATE: 1995-07-20  
PRIOR APPLICATION NUMBER: 08/137,951  
PRIOR FILING DATE: 1993-10-14  
Remaining prior application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 42212  
SOFTWARE: PERL Program  
SEQ ID NO 9510  
LENGTH: 263  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No: hu00416615  
FEATURE:

## Alignment Scores:

Pred. No.:	2,046-18	Length:	311
Score:	334.50	Matches:	93
Percent Similarity:	51.52%	Conservative:	60
Best Local Similarity:	31.31%	Mismatches:	103
Query Match:	9.36%	Indels:	41
DB:	2	Gaps:	11

US-09-513-151-3 (1-2041) x F64046 (1-311)

```

QY 86 GTAAGTATTCGCGGCGCCAGCGCAAAATCCAGCGTGGCTTGACGTAGCCAG 145
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 6 IlePheLeuMetGlyProThrAlaSerGlyThrSpleuAlaIleGlnLeuArgSer 25
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 146 CGGCTGGCGCGGTGATGTCAGCGGTGACTCCATGCGAGGTATGAGGCCCTAGACATC 205
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 26 GlnLeuProValGlnValIleSerValAspSerAlaLeuIleTyrLysGlyMetAspIle 45
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 206 ATGACCAACAAGTTTGTGCGCCAAAGACAGAAATCGCGCGCACACATGATCAGCTTT 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 46 GlyThrAlaLysProSerLysGlnGlnLeuAlaProHisArgLeuIleAspIle 65
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 266 GTGATCCCTCTGTGACCAATTAACAGGTGGACTTCAGAAATAGCAACTGCTCTG 325
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 66 LeuAspPro---SerGlnSerTyrSerAlaMetAsnPheArgAspAlaLeuArgGlu 84
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 326 ATTGAAGATATATTGGCCGAGACAAATTCCTATTGTGTGGAGGAGACCAATTATTAC 385
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 85 MetAlaAspIleThrAlaGlnGlyLysIleProLeuLeuValGlyGlyThrMetLeuTyr 104
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 386 ATTGAATCTCTGCTGTGGAAGTCTGTGCAATACCAAGCCCCGAGAGATGGCAGCTGAG 445
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 105 TyrLysAlaLeu-----IleGlnGlyLeuSerProLeuProSerAlaAspGlu 120
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 446 AAAGTGAATGACGAAAGTGAAGCTTGAA-----AAGGAGATGGCTTTGTA 493
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 121 AsnIle-----ArgAlaGlnLeuGlnGlnLysAlaIleGlnGlnGlyTyrAlaAla 137
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 494 CTTGCACAAAGCCTTAAGCCAGGTGAGACCCGAAATGCGTGCAGAGCTGCATCCACATGAC 553
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 LeuHisThrGlnLeuAlaLysIleAspProIleSerAlaAlaArgIleAsnProSerAsp 157
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 554 AAAGCGAAAGTGGCCAGAGCTTTCAGAGTTCGAGAAACAGGAATCTCTCATAGTGA 613
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 158 SerGlnArgIleAsnArgAlaLeuGlnValPheTyrIleThrGlyLysSerLeuThrGlu 177
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 614 TTCTCCATCGTCACATACGAGAGAGAGTGTGTGTCCTTGGAGGTCTCTGAAGTTC 673
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 178 Leu-----ThrGlnGlnLysGlyGlnAlaLeu-----ProTyrAspPhe 190
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 674 TCTAACCTTCGATCCTTTGGCTTCATGCTGACCGACAGTCTTAGATGAGCGCTTGAT 733
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValIleuHisGlnArgIleGlu 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 734 AAGAGGTGATGACATGCTGTGCTGGCTGTGGAGAACTAGAGATTTTCACAGA 793
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 209 GlnThrPheHisLysMetIleGlnLeuGlyPheGlnAlaGlnValGlnLysLeuTyrAla 228
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 794 CGCTATATATCAGAAATGTTTCGAAATAGCCAGGACTATCAACATGTATCTTCCAA 853
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 Arg---GlyAspLeuAsnIleAsnLeuProSer-----IleArg 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 854 TCAATTTGGCTTCAAGAAATTTACGAGTACCTG----- 886
   :::::||||| ||| :::::||||| ||| :::::||||| |||
Db 241 CysValGlnTyrArgGlnMetTyrPglutyrLeuGlnGlnLysPyrAlaTyrGlnGlnMet 260
   :::::||||| ||| :::::||||| ||| :::::||||| |||
QY 887 ATCAGTACGAGAAATAGCACACTGAGACATGAGTAAACAGCTTCTAAAGAAA 937
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274
   :::::||||| ||| :::::||||| ||| :::::||||| |||

```

Search completed: April 21, 2003, 18:51:09  
 Job time : 56.8883 secs

DB	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200
DB	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200

C; Superfamily: delta(2)-isopentenylpyrophosphate transferase

9500

C:Genetics:  
A:Gene: miaA  
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
3,35e-19	339	98	63	115	40	9
Score:	344.50					
Percent Similarity:	51.11%					
Best Local Similarity:	31.11%					
Query Match:	9.64%					

US-09-513-151-3 (1-2041) x A71473 (1-339)

```

QY 41 CCTGTGGGCGATGGGCTCAGAGGCGCTGACACGACCTACTCTTGTGATGATCTCGG 100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 ProGlnSerPheSerLysMetPheLysArgThr-----ValIleLeuLeuAlaGly 36
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 101 GCCACGGGACCCGCAATCCAGCTGGCGTTGCAGTAGAGCGAGCGGCTGGCGGTAG 160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 ProThrIlySerGlyLysThrAlaValSerLeuLysLeuAlaProLeuValAspGly 56
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 ATCTGACGCGTCACTCCATGACGCTATGAGCGCTAGACATCATCCACCAAGGTT 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 IleIleSerValAspSerMetGlnValTyrGlnGlyMetAspIleGlyThrAlaLysVal 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 TCTGCCACAGACGACAGATCTGCCGACACATGATCAGCTTTGTGATCCTCTGTG 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 77 SerLeuThrAspArgLysGlnValProHisIleLeuLeuAsp---ValCysHisValGln 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 ACCAATTAACAGTGGGAGTCTGAAATAGACCAATGCTGCTGATGAGATATATT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 GluSerPheAsnAlaValAspPheTyrThrHisAlaValGlnAlaCysGlnAspIleLeu 115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 GCCCGAGACAAATCTCTATTGTGTGGGAGAACCAATTAATTAATTAATCTCTGCTC 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 116 SerArgAsnLysValProIleLeuValGlyThrGlyPheTyrPheHisThrPheLeu 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 TGGAAAGTCTTTCATATACCAAGCCCGACAGATGGGACACTGAGAAAGTATTACCGA 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 -----SerGlyProProSerGlyProSerProAspPheValLeuArgLys 150
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 AAATGGAGCTTGAA---AAGGAGATGGTCT---GTATTCACAAACGGCTAGCCAG 514
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 GlnLeuThrLeuGlnAlaGlnGlyLysIleSerAlaLeuTyrGlnGlnLeuLys 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 515 GTGACCCAGAAATGGCTGCCAAGCTGCATCCATGACCAACGCAAGTGGCCAGAGC 574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 LeuAspProValTyrAlaIleThrIleThrLysHisAspLysAsnLysIleIleArgAla 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 575 TTGCAAGTTTGAAGAAACAGAAATCTCTCATAGTGAATTTCTCCATCGTCACATACG 634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 LeuGlnIleIleArgLysThrGlySerLysValSerSerTyrAlaIleTrpGlnSerThrVal 210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 635 GAAAGAGTGGTGGTCCCTGGAGGCTCTGAGAGTCTCAACCCCTGGACCTCTGG 694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 AsnGlnSerLys-----GluTyrHisCysArgArgTrp 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 695 ---CTTCATGCTGACACGACAGTCTAGATGAGCCCTTGATAGAGGGTGGATGACATG 751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 LeuLeuSerProAspProGlnLeuLeuArgHisAsnIleLeuGlnAlaCysAspGlnMet 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 752 CTGCTGCTGGCTCTTGGAGACTAAGAGATTTTTCACAGACGCTATTAATAGAAAGAT 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 LeuGlnGlnGlyLeuLeuAspLysValGlnAlaLeuLeuAlaIleGlyTyrLeuSerLysAsn 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 812 GTTTCGGAATAAGCCAGGACTATCAACATGATGATCTTCATGATGATGATGATGATG 871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 262 SerSerAlaSer-----ArgAlaIleIleGlyTyrArgGln 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 872 TTTCACGAGTACTG-----ATCAGTACGAGGAGAAA 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 TrpIleGlnPheLeuAspLeuGlnLysSerProProAspLeuPheGlnIleThrLysGlnLys 292
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

QY 902 TGCACACTGGAGACTAGTAACCA-GCTTTAAGAAAGAAAGACTG 945  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 293 PheIleThrAsnThrTrpArgTyrThrLysGlnArgThrTrp 307

## RESULT 13

B89905  
 hypothetical protein miaA [imported] - staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: B89905  
 R:Kuroda, M.; Ohka, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.;  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kato, C.; Sekimizu,  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsuku, K.  
 Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant staphylococcus aureus.  
 A:Reference number: A89758; M01D:21311952; PMID:11418146

A:Accession: B89905

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <KUR>

A:Cross-reference: GB:BA000018; PID:g13701103; PIDN:BA842398.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: miaA

C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

## Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
8.96e-19	311	86	62	92	26	7
Score:	339.00					
Percent Similarity:	55.64%					
Best Local Similarity:	32.33%					
Query Match:	9.48%					

US-09-513-151-3 (1-2041) x B89905 (1-311)

```

QY 83 CTGTAGAGATTCCTGGGCGGACCGGACCGCAAAATCCAGCTGGGCTGACGTAGGC 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 IleValValIleValGlyProThrAlaSerLysThrGlnLeuSerIleGlnLeuAla 27
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 143 CAGCGGCTCGGCGGTGACATGCTGACGCTGATCCATGATGATGATGATGATGATGATG 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 LysArgIleAsnGlyGlnIleIleSerGlyAspSerMetGlnValTyrLysHisMetAsn 47
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 203 ATCATCAACCAAGTCTTGGCCCAAGACGACAGAAATCTCGCGCACCAATGATGACG 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 48 IleGlyThrAlaLysValThrProGlnGlnMetAspGlyIleProHisIleLeuLeuAsp 67
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 263 TTTGGATCGCTTGTGACCAATTAACACAGTGGGACCTGACAAATAGACCAACTGCT 322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 IleLeuAsnProAspAspThr---PheSerAlaTyrGlnPheLysArgLeuAlaGlnAsp 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 323 CTGATTTGAAGATATTTTGGCCGACACAAATCTCTATTGTGTGGGAGAACCAATTAT 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 87 LeuIleThrAspIleThrAsnArgGlyLysValProIleIleAlaGlyThrGlyLeu 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 TACATTTGATCTCTGCTGCGAAAGTCTTGTCAATFACCAAGCCCGACAGATGGCGACT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 107 TyrIleGlnSerLeuIleTyrAsnTyrGlnLeuGlnAspGlnThrValThrProAlaGln 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 GAGAAAGTATGATGACCGAAAGAGTGAG---CTTGAAGAGAGAGATGCTTGTACTTAC 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 LeuSerIleValLysGlnLysLeuSerAlaLeuGlnHisLeuAspAsnGlnGlnLeuHis 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 500 AAACGCTTAAGCCAGGCTGGACCCAGAAATGGCTGCCAGCTGCATCCATGACAAACGC 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 AspTyrLeuAlaGlnPheAspAlaValSerAlaGlnAsnIleHisProAsnAsnArgGln 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 560 AAAGTGGCCAGGAGCTTGCACATTTTGAAGAAACAGG-----ATCTTCATAGTAA 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 167 ArgValLeuArgAlaIleGlnTyrTyrLeuLysThrLysLysLeuLeuSerAsnArgLys 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db 103 TyrIleasnSerIleuIecysaenTyrAspPheThrGlyAlaTyrIysaspGluAlaTyr 122
QY 440 ACTGAGAAAGTATTGACCGAAAAGTGAGCTTGAAGAGAGAGATGCTT---GTACTT 496
Db 123 ArgGluSerLeu-----GlnAlaIleAlaIysaspGlyLysGluTyrLeu 138
QY 497 CACAAAGGCTTAAGCCAGGTGACCCAGAAATGGCTGCCAAGCTGCATGCATGACAA 556
Db 139 HisGluLysLeuLysasnIleAspSerTyrLysIleuTyrProAsnAspLeu 158
QY 557 CGCAAGTGGCCAGAGCTTGCAGTTTGAAGAACAAGCATCTCATGTGATTT 616
Db 159 LysArgValIleArgAlaLeuGluValTyrLysIleThrGlyLysThrIleSerGluLeu 178
QY 617 CTCATCGTCACATACAGAGAGAGAGGTGGTCCCTTGAGAGTCTCTGAATTCCT 676
Db 179 -----AsnSerAsnValAspLeuTyrAspIleProTyr----- 189
QY 677 AACCTGTCATCTTGGCTTCACTGCTGACACGAGCTTTCAGATGAGCGCTGATAG 736
Db 190 AsnIleHisTyrPheIleLeuAsnMetAspArgGlnLysLeuTyrGluArgIleAsnLeu 209
QY 737 AGGCTGATGACATGCTGCTGCTGCTGCTGCTGAGAGAA-----CTAGAGATTTT 787
Db 210 ArgValAspIleMetLeuArgAsnGlyLeuValAspGluValIleLysLeuArgAspMet 229
QY 788 CACAGAGCTTAATACAGAGATGTTTCGAAAATAGCCAGACTATACATGATGATC 847
Db 230 GlyTyrAsnSerAsnMetClnSerMet----- 238
QY 848 TTCATATCATTTGGCTTCAAGGATTTACAGATCTGATACGAGGAAATGACACA 907
Db 239 ---LysGlyIleGlyTyrLysGlnIleuSerTyrLeu-----GlnGlyCysIleThr 255
QY 908 CTGAGACTAGTAACACGCTTCTAAAGAAAGA 940
Db 256 LeuGluGluAlaValGlnIleuIleLysGly 266

RESULT 11
C72366
tRNA delta-2-isopentenylpyrophosphate transferase - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: C72366
R:Neelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: C72366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <ARN>
A:Cross-references: GB:AE001728; GB:AE000512; NID:94981027; PIDN:AA035610.1; PID:9498103
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0525
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Alignment Scores:
Pred. No.: 1,72e-19 Length: 305
Score: 348.00 Matches: 86
Percent Similarity: 53.66% Conservative: 68
Best Local Similarity: 29.97% Mismatches: 105
Query Match: 9.73% Indels: 28
DB: 2 Gaps: 8

US-09-513-151-3 (1-2041) x C72366 (1-305)
QY 77 CTACCTCTGTAGTATTCTCGGCGACGACGACGACGACGACGACGACGACGACGACG 136
Db 1 MetLysIleAlaIleValGlyLysProThrAlaValAlaGlyLysThrAspIleMetIleGlu 20

```

```

QY 137 CTAGCCACGCGCTCGCGGTGATGATGTCAGCGCTGACTCCATGCGAGCTATGAAAGC 196
Db 21 ValCysGluGluIleGlyAlaIleGluIleIleSerMetAspSerArgIleIleTyrArgTyr 40
QY 197 CTAGCATCATCATCACCAAGAGTTTGTGCCCAAGACAGACAGAGATTCGCCGACACAG 256
Db 41 MetAspIleGlyThrAlaLysProThrProGluGlnArgLysArgValIleuHisIleMet 60
QY 257 ATCAGCTTGTGATGATCTTGTGACCAATTCACAGCTGGTGGACTTCGCAATATACGA 316
Db 61 IleAspIleIleSerPro---AspGluTyrTyrAsnAlaPheMetTyrArgLysAspSer 79
QY 317 ACTGCTCTGATTAAGATATATTTGCGCCAGACAAATTCATATGTTGTGGAGAGAC 376
Db 80 LeuArgAlaMetClnAspAlaLeuArgAlaGlyLysIleProValTyrAlaGlyGlyThr 99
QY 377 AATTATTCATGATGATCTGCTCTGGAAGATCTTGTCAATACCAAGCCCAAGAGATG 436
Db 100 GlyLeuTyrAlaAspAlaLeuVal---ArgGlyIlePheGluGlyAlaProAlaAspGlu 118
QY 437 GGCACCTGAGAAAGTGTATACCGGAAAGTGAGAGCTTGAAGAGAGATGCTTACTT 496
Db 119 AsnIleArgLysGluLeuArg-----GluLeuGlnArgArgGluProGlyIleLeu 135
QY 497 CACAAAGGCTTAAGCCAGAGTGGACCCAGAAATGGCTGCCACAGCTGATCCACATGACAA 556
Db 136 ArgLysMetLeuGluGluIleuAsnProGluAlaAlaThrArgIleHisProAsnAspLeu 155
QY 557 CGCAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAAATCTCATAGTGAATTT 616
Db 156 LysArgThrIleArgAlaLeuGluValTyrMetLysThrGlyArgArgIleSerGluLeu 175
QY 617 CTCATCGTCACATACAGAGAGTGGTGGTCCCTTGAGAGTCTCTGAATTCCT 676
Db 176 -----GlnLysGluAlaLysGlyAspAsp-----ArgPhe--- 185
QY 677 AACCTGTCATCTTGGCTTCACTGCTGACACGAGCTTTCAGATGAGCGCTGATAG 736
Db 186 -----PheIleIleValLeuThrArgGluArgTyrGluLeuTyrGluArgIleAsnLys 203
QY 737 AGGCTGATGATGATGCTTCTGCTGCTGCTGCTGAGAGAACTAAGCATTTTACAGAGCC 796
Db 204 ArgValAspLysMetIleGluMetGlyLeuValAspGluValLysArgLeuGlyMet 223
QY 797 TATATCAACAAGATGTTGCGAAATATACCAAGACTATCAACATGATCTTCCAATCA 856
Db 224 GlyTyrSerLysAspLeuAsnSer-----MetLysThr 234
QY 857 ATTGCTTCACAGAAATTTCCAGAGTACATGATCAGAGGAAATGCACACTGAGACT 916
Db 235 IleGlyTyrLysGluValIleAspTyrLeu-----GluGlyLysTyrAspPheAspLys 252
QY 917 ACTAACACGCTTCTTAAGAA 937
Db 253 MetValHisLeuIleLysArg 259

RESULT 12
A71473
probable tRNA pyrophosphate transferase - Chlamydia trachomatis (serotype D, strain
C:Species: Chlamydia trachomatis
C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: A71473
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mltc
Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia
A:Reference number: A71570; MUID:9900809; PMID:9784136
A:Accession: A71473
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-339 <ARN>
A:Cross-references: GB:AE001349; GB:AE001273; NID:93322226; PIDN:AA068361.1; PID:93
A:Experimental source: serotype D, strain UW-3/Cx

```



```

|||||  |||  |||:|||||  |||||  |||||:|||||
47 AsplIleGlyThrAlaLysIleThrGluGlnGluMetGluGlyValProHisHisLeuIle 66
QY 260 ACCTTGGTGCATCTCTGTGTGACCAATTCACAGTGGTGACTTCAGAAATGAGCACT 319
Db 67 AsplIleuAspPro---GlnAspSerPheSerThrAlaAspTylGlnSerLeuValArg 85
QY 320 GCTGATGAGATATATTTTCCCGAGACAAATTCATATGTTGTGTGGAGAACCAAT 379
Db 86 AsnLysIleSerGluIleLalaSnaIrgLysLeuProMetCileAspLysIleThrGly 105
QY 380 TATTACATTAATCTCTGCTGTGAAAGTTCTTCATATACCAAGCCCGAGAGATGGGC 439
Db 106 LeuTyrIleGlnSerGluLeuTyrAspTyrThrPheThr-----GluGluAla 121
QY 440 ACTGAGAAAGTATGACCCGAAAGTGAGCTT-----GAAAGAGAGATGCTTTGTA 493
Db 122 AsnaSproValPheArgLysSerMetGlnMetAlaIaIaIaIaIaIaIaIaAspPhe 141
QY 494 CTTCACAAAGCCGTAAGCCAGGTGACCCGAAATGGCCGCAAGCTGCATCCACATGAC 553
Db 142 LeuHISAlaLysLeuAlaIaIaIaAspProGluAlaIaIaIaIaIaIaIaIaIaIaIa 161
QY 554 AAACGCAAGTGGCCAGAGCTTGCAGATTGTAAGAAACAGCAATCTCATAGTGA 613
Db 162 ThrArgArgValIleArgAlaLeuGluIleLeuHISThrSerGlyLysThrMetSerGln 181
QY 614 TTTTCCATCGTCACATACAGAGAAGGTGTGTCCTCCCTTGAGAGTCTGTGAATTC 673
Db 182 HISLeuLysGluGlnLysArgGlu-----LeuLeu 191
QY 674 TCTAACCCCTGCATCCTTTGGCTTCATGCTGACAGAGCACTTCATAGAGCGCTTGAT 733
Db 192 TyrAlaAlaValLeuIleLysLeuThrMetAspArgAspThrLeuTyrGluIleAsn 211
QY 734 AAGAGGCTGATGATACATGCTTGTGCTGGCTCTTGAGAGCACTAAGATTTTCACA 793
Db 212 GlnArgValAspLeuMetMetGlnSerGlyLeuLeuProGluVal-----Lys 227
QY 794 CCCTAATACAGAGATGTTTCGAAATATGACAGACTATCACATGATCTTCCAA 853
Db 228 ArgLeuTyrAspLysAsnVal-----ArgAspCysGlnSer-----IleGln 241
QY 854 TCAATTGGCTTCAGGAATTTTCACAGATCCTGCATCAGTACGAGGAGAAATGCACACTGGAG 913
Db 242 AlaIleGlyTyrLysGlnLeuTyrAlaTyrPhe-----AspGlyPheValThrLeuSer 259
QY 914 ACTAGTACCACTTCTAAGAA 937
Db 260 AspAlaValGluGlnLeuLysGln 267

RESULT 8
AF1236
tRNA Isopentenylpyrophosphate transferase homolog miaa [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1236
R:Glasner, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.;
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitouram, A.; Ma
O, C.; Schlueter, T.; Simoes, N.; Tlherre, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-305 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CA093972.1; PID:gl6410710; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: miaa

```

```

C:Superfamily: delta(2)-Isopentenylpyrophosphate transferase
Alignment Scores:
Pred. No.: 2e-23 Length: 305
Score: 397.50 Matches: 98
Percent Similarity: 54.51% Conservative: 59
Best Local Similarity: 34.03% Mismatches: 100
Query Match: 11.12% Indels: 31
Db: 2 Gaps: 7
US-09-513-151-3 (1-2041) x AF1236 (1-305)
QY 77 CTACCTCTGTGATGATGATGCGGGGACAGCGGCAATTCACAGCTGGCTTGAG 136
Db 4 IleProValIleValIleValIleValIleProThrAlaValGlyLysThrSerLeuSerIleGlu 23
QY 137 CTAGCCAGCGGCTGCGCGGTGAGATCGTCAGCGGTGATCCATCCAGTCTATGAAAGC 196
Db 24 LeuAlaLysLysLeuAspLysGluIleIleSerGlyAspSerMetGlnValTyrArgLys 43
QY 197 CTAGACATCATCACACAGAGTTTCTGCCACAGACACAGATATCGCCGACACATG 256
Db 44 LeuAspIleGlyThrAlaLysIleThrProGluGluMetAspGluIleLysHisTyrLeu 63
QY 257 ATCAGCTTGTGATCCCTCTTGACCAATTCACAGCTGCTGACTTCAGAAATGAGCA 316
Db 64 IleAspValThrAspPro---SerGluProPheThrAlaAlaLysPheIleThrGluThr 82
QY 317 ACTGCTGTGATGAGATATATTTCGCCGAGCAAAATTCATATGTTGTGGAGAAC 376
Db 83 ArgLysThrPheIleGluThrIleHisGlnAlaGlyLysLeuProIleLeuAlaGlyLysThr 102
QY 377 AATTATTCATGAAATCTCTCTGCGAAAGTTCTTGCAATACCAAGCCGAGAGATG 436
Db 103 GlyLeuTyrIleGlnSerValPheTyrAspTyrAspPheGlyAsnValSerGluAspLys 122
QY 437 GGCCTAGCAAGATGATTCACCGAAAGTGAGCTTGAAAGAGAGATGCTTGTACT 496
Db 123 Ala-----TyrArgAlaGluLeuGlnGluLeuAsnLysThrLeu 136
QY 497 CACAAAGCCCTAAGCAGGTGAGCCAGAAATGGCTGCCACATCATCAGATACAA 556
Db 137 TrpGlnMetLeuGlnGlnGlnAspProGluSerAlaIaIaGlnIleHisGlnAsnLys 156
QY 557 GCCAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGGAATCTTCATATGAAATT 616
Db 157 ArgArgValIleArgAlaLeuGlnValMetHisLeuThrGlyLysProPheSerGluTyr 176
QY 617 ---CTCCATCGTCACATCAGCAAGAGGTGTGTCCTCCCTTGAGAGTCTCTGAAGTTC 673
Db 177 GlnValHisAsnValLeuAsnAspThr----- 185
QY 674 TCTAACCCCTGCATCCTTTGGCTTCATGCTGACAGCAAGCTTCATAGAGCGCTTGAT 733
Db 186 TyrLysProLeuPheLeuGluLysLeuAspLeuAspArgAlaLeuLeuTyrGluIleAsn 205
QY 734 AAGAGGCTGATGATACATGCTTCTGCTGGCTTGGAGAGAACTAAGATTTTCACA 793
Db 206 GlnArgValAspLeuMetPheGlnGluGlyLeuValThrGluAlaLys-----Lys 222
QY 794 CCCTAATACAGAGATGTTTCGAAATATGACAGACTATCACATGATCTTCCAA 853
Db 223 LeuTyrAspGlnHisLeuAlaAspValProAla-----ValCys 235
QY 854 TCAATTGGCTTCAGGAATTTTCACAGATCCTGCATCAGTACGAGGAAATGCACACTGGAG 913
Db 236 GlyIleGlyTyrLysGlnLeuPheProTyrPhe-----GluGlyLysSerSerLeuGlu 253
QY 914 ACTAGTACCACTTCTAAGAA 937
Db 254 GlnAlaLysGluLeuIleGlnLys 261

RESULT 9

```

QY 1193 -----AACCACTGAAGAAAGAGAGATGTCAGTCTGCTCAAC 1237  
 DB 442 ThirgH1sLysAsnSerGlnThrTyLysAsnArgLValGlnGlnValAsn 461  
 RESULT 6  
 F83945  
 tRNA Isopentenylpyrophosphate transferase miaa [Imported] - Bacillus halodurans (strain  
 C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
 C:Accession: F83945  
 R:Takekmi, H.; Nakasone, K.; Takeki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
 A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: F83945  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-314 <STO>  
 A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA06085.1; GSPDB:GN00  
 A:Experimental source: strain C-125  
 C:Genetics:  
 A:Gene: miaa  
 C:Superfamily: delta(2)-isopentenylpyrophosphate transferase  
 Alignment Scores:  
 Pred. No.: 5,67e-25 Length: 314  
 Score: 417.00 Matches: 102  
 Percent Similarity: 56.27% Conservative: 64  
 Best Local Similarity: 34.58 Mismatches: 85  
 Query Match: 11,664 Indels: 44  
 DB: 2 Gaps: 10  
 US-09-513-151-3 (1-2041) x F83945 (1-314)  
 QY 83 CTGTGATGATCTTCGGGGCGGCGACCGCAATCCAGCTGGCGTTCAGCTAGGC 142  
 DB 5 LeuValAlaIleValGlyProThrAlaValAlaGlyLysThrLysThrSerValMetLeuAla 24  
 QY 143 CAGCGGCTCGGCGGTGATGATGTCAGCGCTGCTCCATGAGCTATGAGCGCTTAC 202  
 DB 25 LysArgLeuAsnGlyGluValIleSerGlyAspSerMetGlnValTyrArgGlyMetasp 44  
 QY 203 ATCATGACCAAGAGTTCTGCGCCAGACAGAGATCTGCGGCGCCACATGATAGC 262  
 DB 45 IleGlyThrAlaLysIleThrAlaGluGluMetaspGlyValProHishLysLeuIleasp 64  
 QY 263 TTGTGTGATCTTCGTGACCAATACACAGCTGAGCTTCAAGAAATAGCAACTGCT 322  
 DB 65 IleLysAspPro---SerGluSerPheSerValAlaAspPheGlnAspLeuAlaThrPro 83  
 QY 323 CTGATGAGAGATATATTTGGCCGACAGCAAAATCTTATGTTGTGGAGGACCAATAT 382  
 DB 84 LeuIleThrGluIleHsGlnArgGlyArgLeuProPheLeuValGlyGlyThrGlyLeu 103  
 QY 383 TACATGAACTCTGCTCTGGAAGT---CTGTCAATATCCAGCCCGGAGATGGC 439  
 DB 104 TyrValAsnAlaValIleHsGlnPheAsnLeuGlyAspIleArgAlaAspGlu----- 121  
 QY 440 ACTGAGAAAGTGAATGACGAAAGTGAGACTTGA-----AAGGAGATGCTCT 490  
 DB 122 -----AspTyrArgHsGlnLeuGlnAlaPheValAsnSerTyrGlyVal 136  
 QY 491 ---GTACTTCACAAGCGCTTAACCCAGGTGAGCCAGCAAAATGCTGCGCAAGTGCATCA 547  
 DB 137 GlnAlaIleLysAspLysLeuSerTyrIleAspProLysAlaAlaIleHsPro 156  
 QY 548 CAGCACAAGCGCAAGTGGCGGAGAGCTTCAAGTTTAAAGAAAGCAATCTCAT 607  
 DB 157 AsnAspTyrArgValIleArgAlaLeuGlnIleLysLeuThrGlyLysThrVal 176  
 QY 608 AGGAATTTCTCATGCTCAACATAGGAGAAAGTGTGCTCCCTTGAAGGCTCTCTG 667  
 DB 667

DB 177 ThrGlu-----GlnAlaArgHsGlnGluGlnThrProSerProTyr----- 190  
 QY 668 AAGTCTTAACCCCTTCATCTTTGGCTTCATGCTGACCGCACTTCTAGATGAGCC 727  
 DB 191 -----AsnLeuValMetIleGlyLeuThrMetGlnArgAspValLeuTyrAspArg 207  
 QY 728 TTGCATAGAGGCTGATGATGATGCTGCTGCGGCTCTGGAGAGCAATGATTT 787  
 DB 208 IleAsnArgArgValAspGlnMetValGluGlnGlyLeuIleAspGlnAlaLysLysLeu 227  
 QY 788 CACAGACGCTATATATACAGAAATCTTTCGAAATACCGACGATCATCATGATATC 847  
 DB 228 -----TyrAspArgGlyIle 232  
 QY 848 -----TTCCATCAATGCTTTCAGGAATTCACAGTACGATGATCACT 892  
 DB 233 ArgAspCysGlnSerValGlnAlaIleGlyTyrLysGlnMetLysAspTyrLeu----- 250  
 QY 893 GAGGCAAAATGACGATGAGTACCAACCACTTCTTAAGAA 937  
 DB 251 AspGlyAsnValThrLeuGlnAlaIleAspThrLeuLysArg 265  
 RESULT 7  
 G69657  
 tRNA Isopentenylpyrophosphate transferase miaa - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
 C:Accession: G69657  
 R:Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; B  
 C:Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, A.B.; Capuano, V.; Carter, N.M.;  
 A:Erlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga  
 lech, J.; Harwood, C.R.; Henault, A.; Hilbert, H.; Holstappel, S.; Hosono, S.; Hullo,  
 Koester, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardin  
 Y: M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scan  
 A:Authors: Schlecht, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowsky, A.; S  
 akeuchi, M.; Yamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchly  
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshic  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: G69657  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-314 <KUN>  
 A:Cross-references: GB:299113; GB:AL009126; NID:g2634090; PIDN:CA013617.1; PID:g2634  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: miaa  
 C:Superfamily: delta(2)-isopentenylpyrophosphate transferase  
 Alignment Scores:  
 Pred. No.: 6,71e-24 Length: 314  
 Score: 403.50 Matches: 97  
 Percent Similarity: 55.56% Conservative: 63  
 Best Local Similarity: 33.68 Mismatches: 99  
 Query Match: 11,29% Indels: 29  
 DB: 2 Gaps: 8  
 US-09-513-151-3 (1-2041) x G69657 (1-314)  
 QY 80 CTTCTGTAGTATCTTCGGGCGGCGACCGCAATTCACGCTGGCTGACGCTA 139  
 DB 7 ProValValIleLeuValGlyProThrAlaValAlaGlyLysThrAsnLeuSerIleGlnLeu 26  
 QY 140 GCGCAGCGGCTCGGCGGTGATGATGTCAGCGCTGATGATGAGCTATGAGCGCTA 199  
 DB 27 AlaLysSerLeuAsnAlaGlnIleIleSerGlyAspSerMetGlnIleTyrLysGlyMet 46  
 QY 200 GACATCATCAACAAGATTTCTGCCCAAGACAGAGATCTGCGGCGACCAATGATC 259

OY 989 -----TCGACGAGGAGGAGCTGTTCTGTAACCTGCTTGAATGCTGCA 1036  
 DB 371 SerLysSerGluGluSerTrpAsnAlaGlnValAlaValysProAlaSerGluIleLeuArg 390  
 OY 1037 AGTTTCATCCAG-----GGCCACAAGCCTACAGCCATCCATAAAGATG 1081  
 DB 391 CysPheLeuGluThrGluThrGluSerGlyArgAspProThrSerGlyLys----- 407  
 OY 1082 CCATACAAATGAGCTGAGACAGACAGAAAGATTATCATCCTGTGACCTGTG--GATGCA 1138  
 DB 408 -----SerIleGluArgAspLeuTrpThrGlnTrpValGlyGluAlaCysGlyAsnLys 425  
 OY 1139 ATCATCATTTGGGGATGGCGATGGCCAGCCATAAATCCAAATCCCACTTG----- 1192  
 DB 426 IleLeuArgGlyArgHisGluTrpGluHisLysGlnGlyArgThrHisArgLysArg 445  
 OY 1193 -----AACCAACTGAAGAAAGAGAGAGAGATGTGAGCTCAGATCGCTGC 1234  
 DB 446 ThrTrpArgHisLysAsnSerGlnThrTyrrLysAsnArgGluValGlnGluAlaGluVal 465  
 OY 1235 AAC 1237  
 DB 466 Asn 466

RESULT 5  
 F84676  
 hypothetical protein At2g27760 (imported) - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_rev:1001 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: F84676  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Unayam, L.; Tallon, L.;  
 Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: F84676  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-461 <STO>  
 A:Cross-references: GB:AE002093; NID:93860256; PID:NAC73024.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g27760  
 A:Map position: 2

Alignment Scores:  
 Pred. No.: 3,2e-29 Length: 461  
 Score: 471.00 Matches: 128  
 Percent Similarity: 48.96% Conservative: 107  
 Best Local Similarity: 26.67% Mismatches: 137  
 Query Match: 13.17% Indels: 108  
 DB: 2 Gaps: 17

US-09-513-151-3 (1-2041) x F84676 (1-461)

OY 65 CTGACAGGACCTACCTCTGTAGTATTCGGGGCAGCGGACGACGCAATCCACG 124  
 DB 1 MetLysLysAlaValValValIleMetGlyProThrGlySerGlyLysSerLys 20  
 OY 125 CTGGGTTGACCTAGCGGACGCGCTCGCGGTGAGATCGTACGGCTGACTCCATGAG 184  
 DB 21 LeuAlaValAspLeuAlaSerHisPheProValGluIleIleAsnAlaAspAlaMetGln 40  
 OY 185 GTCTATGAGGCTGATCATCATCACAACAAGTTTCTGCCACAGACAGAGATCGC 244  
 DB 41 IleTrpSerGlyLeuAspValLeuThrAsnLysValThrValAspGlnGlnLysGlyVal 60  
 OY 245 CGGACACATGATCAGCTTTGTGATCCTCTTGACCAATTACACAGTGGTGACTTC 304  
 DB 61 ProHisHisLeuLeuGlyThrValSerSerAspMet---GluPheThrAlaArgAspPhe 79  
 OY 305 AGAATATAGACACTGCTCTGATTGAGATATATTGGCCGAGACAAATTCATTGTT 364

DB 80 ArgAspPheThrValProLeuIleGluIleValaSerArgAsnHisIleProValLeu 99  
 OY 365 GTGGGAGGACCAATTTATCATGATGATCTGCTGGAAGTTCTTGCAATACCAAG 424  
 DB 100 ValGlyGlyThrHisIleTyrIleGlnAlaValValSerLysPheLeuAsnAspAla 119  
 OY 425 CCCCGAGATGGGACCTGAGAA-----GTATGACCGGAAA 463  
 DB 120 AlaGluAsp-----ThrGluGluCysCysAlaAspValAlaSerValValAspIleAsp 137  
 OY 464 GTGGAGCTTGA-----AAGAGAGATGCTCTTGACTCATTCAACAGGCTTACG 511  
 DB 138 MetValValGluSerValPheGlyArgAspAspLeuSerHisGlyTyrGluLeuLys 157  
 OY 512 CAGGTGACCCAGAAATGCTGCCAAGCTGCATCCATGACACAAAGCAAGTGGCCAG 571  
 DB 158 GluLeuAspProValAlaAlaAsnArgIleHisProAsnAsnHisArgLysIleGln 177  
 OY 572 AGCTTGCAAGTTTGAAGAAACAGAAATCTCTCATGATTTCTGCATCGTACAT 631  
 DB 178 TyrLeuSerLeuHisAlaSerArgGlyValLeuProSerLysLeuTyrrGlnGlyThr 197  
 OY 632 ACGGAA-----GAAGGTGGTGGTCCCTTGAGAGTCTCT 664  
 DB 198 AlaGluValAspSerPhePheLeuThrSerArgGlnAsnTrpGlyCysIleAsnAlaSer 217  
 OY 665 CTGAGTTCTCTAACCTTGCATCCTTGGCTTCAAGCTGACAGCAGATCTAGATGAG 724  
 DB 218 ---ArgPhe---AspTyrCysLeuIleCysMetAspAlaGlnThrAlaValLeuAspArg 235  
 OY 725 CGCTTGATTAAGAGGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784  
 DB 236 TyrValGlnGlnArgValAlaAspAlaMetValAspAlaGlyLeuLeuAspGluValLysAsp 255  
 OY 785 TTTCACAGACGCTATTAACAGAAAGATTTGCGAAATAGCCAGACATCAACATGCT 844  
 DB 256 IleTyrrLys-----ProGlyAlaAspTyrrThrArgGly 266  
 OY 845 ATCTTCACATCAATGAGCTTCAAGAAATTCACAGTACTG-----ATCAGTACG 895  
 DB 267 LeuArgGlnSerIleGlyValAlaArgLysPheGlnAspPheLeuLysIleHisLeuSerGln 286  
 OY 896 -----GGAATATGACACA---CTGGAGACTACTAACACAGCTTCTAAAGACGACT 943  
 DB 287 ThrCysAlaGlnHisLeuThrSerLeuSerAsnAspLysValIleMetLysGlnAsnLeu 306  
 OY 944 GGTCCCAT----- 952  
 DB 307 ArgLysIleLeuAsnPheProLysAspAspLysLeuArgIleMetLeuGlnGluAlaIle 326  
 OY 952 ----- 952  
 DB 327 AspArgValLysLeuAsnThrArgArgLeuLeuAlaArgGlnLysArgValSerArg 346  
 OY 953 GTCCCCCTGTCTATGCTTACAGGTATCTGATGTC----- 988  
 DB 347 LeuGluThrValPheGlyTyrAsnIleHisTyrIleAspAlaThrGluTyrIleLeuSer 366  
 OY 989 -----TCGAATGGGAGAGAGCTGTCTTGAACCTGCTCTGGAATAGCTGCAAGCT 1039  
 DB 367 LysSerGluGluSerTrpAsnAlaGlnValAlaValysProAlaSerGluIleLeuArgCys 386  
 OY 1040 TTCATCCAG-----GGCCACAAGCCTACAGCCATCCATAAAGATGCA 1084  
 DB 387 PheLeuGluThrGluThrGluSerGlyArgAspProThrSerGlyLys----- 402  
 OY 1085 TACAATGAGCTGAGACAGACAGAAAGATTATCATCCTGTGACCTGTG--GATGCAATC 1141  
 DB 403 ---SerIleGluArgAspLeuTrpThrGlnTrpValGlyGluAlaCysGlyLysAsnLysIle 421  
 OY 1142 ATCATTTGGGAGTGGGATGGCGATGGCCAGCCATAAATCCCACTTG----- 1192  
 DB 422 LeuArgGlyArgHisGluTrpGluHisLysGlnGlyArgGlyThrHisArgLysArgThr 441

```

Db 210 AspSerLeuValLeuMetProArgLeuAspLysArgValAspLysMetLeuSerHisGly 229
QY 764 CTCCTGGAGGAACTAAGAGATTTCACAGACCTAATACAGAGATGTTCCGAAAT 823
Db 230 LeuValAspGluIleLys-----SerMetLysSerLeuAlaGluSer 243
QY 824 -----AGCAGAGATCATCAATGATCTTCATCAATTCATGCTTCAGAAAT 874
Db 244 GluLysPheSerProAspPheThrArgGlyIleTyrPheLysIleLysGluPhe 263
QY 875 CACGAGTACCTG-----ATCAGTACGAGAAATGCACACTGAG----- 913
Db 264 MetProTyrPheGluAlaProSerAspIleValAlaPheAsnAspCys---LeuGluArgMet 282
QY 914 -----ACTAGTACCAAGCTTTAAAGAA----- 937
Db 283 LysValSerThrArgIleTyrAlaLysSerGluLysTyrIleGlnSerArgPheLeu 302
QY 938 -----GGACCTGGTCCATTGTCCTCCCTGCTGCTATGGC 970
Db 303 PrometCysLeuAlaGlnIleAspLeuSerProSerSerIleLeu-----PheSerThr 320
QY 971 TTAGAGATCTGATGCTCGAAGTGGAGAGAGTCTGTTCTGAACTGCTCTTGAATC 1030
Db 321 ThrAsnThrThrAspLeuAsnAsnTyrGluGlnVal---GluLysAlaCysArgVal 339
QY 1031 GTGCAAAAGTTTCATCCAGGCGCCACAAAGCTACAGCCATCCATTAAGATGCATCAAT 1090
Db 340 PheGlnTyrPhePheTyrAsnGlyAspAlaIleAla---ProSerAlaAspAspGlnHis 358
QY 1091 GAAGCTGAGACAGACAGAGAGTTAT----- 1114
Db 359 AlaPheGluLysAlaArgAspTyrLeuSerIleMetAsnGlyArgGlnSerGluLys 378
QY 1115 ---CACCTGTGACCTCTGT---GATCGA-----ATCATCATTTGGGGAT 1153
Db 379 LysPheValCysGluGluLysLeuAspLysArgGlyAspProPheThrValIleGlyGlu 398
QY 1154 CCGGAATGGGACGACGACATTAATCCCAATCCCACTTGACCACTGAAAGAAAGAGA 1213
Db 399 AspAlaPheAsnValHisIleLysSerArgLysHisLysThrThrValArgArgLysLys 418

RESULT 4
T52061
tRNA Isopentenyltransferase (EC 2.5.1.8) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52061
R:Golovko, A.; Hjaltn, G.
submitted to the EMBL Data Library, November 1998
A:Description: A tRNA Isopentenyl transferase from Arabidopsis thaliana.
A:Reference number: Z25928
A:Accession: T52061
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-466 <GO>
A:Cross-references: EMBL:AF109376; PIDN:AAF00582.1
C:Keywords: transferase

Alignment Scores:
Pred. No.: 5,72e-31 Length: 466
Score: 493.00 Matches: 133
Percent Similarity: 50.108 Conservative: 108
Best Local Similarity: 27.658 Mismatches: 138
Query Match: 13.798 Indels: 102
DB: 2 Gaps: 18

US-09-513-151-3 (1-2041) x T52061 (1-466)
QY 41 CCTGTGGGCGAGTGGCTCAGAGGC-----CTGCACGCGACCTCTTGTGATGATT 94
Db 6 ProSerAsnGluLysIleGluGluLysMetLysLysLysAlaLysValValIle 25

```

```

QY 95 CTCGGGCGCCAGGCGACCGCAAAATCCAGCTGGCTTGCAGCTAGCGCCGCTCGGC 154
Db 26 MetGlyProThrGlySerGlyLysSerLysLeuAlaValAspLeuAlaSerHisPhePro 45
QY 155 GGTGAGATGTCAGCGCTGACTTCATGAGCTGTATGAGGCTTACATCACCAC 214
Db 46 ValGluIleIleAsnAlaAspAlaMetGlnIleTyrSerGlyLeuAspValLeuThrAsn 65
QY 215 AAGCTTTCGCCAAGAGAGAAATCTCCCGGCGCCACCATATGATACGTTTGGATCT 274
Db 66 LysValThrValAspGluGlnLysGlyAlaProHisHisLeuMetGlyThrValSerSer 85
QY 275 CTGTGACCAATTCACAGCTGCTGCTCAGAAATAGACAACAGTCTGTGATGAAGT 334
Db 86 AspMet---GluPheThrAlaArgAspPheAlaGAspPheThrValProLeuIleGluGlu 104
QY 335 AATATTTGCCGAGACAAATTCCTATGTTGTTGGAGAGAACCAATATTATTCATGAATCT 394
Db 105 IleValSerArgAsnHisIleProValIleGlyGlyThrHisTyrTyrIleGlnAla 124
QY 395 CTGCTGTGAAAGTTTGTTCATACCAAGCCCGGAGATGGGCACTGAGAA----- 448
Db 125 ValValSerLysPheLeuAspAspAlaAlaGluAsp---ThrGluLysCys 142
QY 449 -----GTGATTCACGAGAAAGTGGACCTGGA-----AAGAG 481
Db 143 AlaAspValAlaSerValAlaAspGlnAspMetValValGlnSerValPheGlyArgAsp 162
QY 482 GATGCTCTTGTACTTCACAAAGCCCTTACGCGAGTGGAGCCAGAAATGGCTGCCAGCTG 541
Db 163 AspLeuSerHisGlyTyrIleLeuLeuLysGluLeuAspProValAlaAlaAsnArgIle 182
QY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTCCAAAGTTTGAACAAACAGGAATC 601
Db 183 HisProAsnAsnHisArgLysIleAsnGlnTyrLeuSerLeuHisAlaSerArgGlyVal 202
QY 602 TTCATATGAAATTTCTCCATGCTCAACATGAGGAAGAGTGGTCCCTGGAGAGT 661
Db 203 LeuProSerLys---LeuTyrGlnGlyLysThrAlaGlnAsnTyrGlyLysIleAsnAla 221
QY 662 CCTGTGAAGTTCTTAACCTTCGATCCCTTGGCTTCATGCTGACGAGCAGCTAGTAT 721
Db 222 Ser---ArgPhe---AspTyrCysLeuIleIleCysMetAlaGluThrAlaValLeuAsp 239
QY 722 GAGCCCTGGATTAAGAGGTGGATGACATGCTGCTGCTGGCTTGGAGAGACTAATA 781
Db 240 ArgTyrValGluGlnArgValAlaAspAlaMetValAspAlaGlyLeuLeuAspGluValTyr 259
QY 782 GATTTCACAGACGCTATATCAGAAAGATGTTTCGAAATATGACGAGCATCAACAT 841
Db 260 AspIleTyrLys-----ProGlyAlaAspTyrThrArg 270
QY 842 GGTATCTTCATCAATTCGCTTCAAGAAATTCACGAGTACCTG-----ATCACT 892
Db 271 GlyLeuArgGlnSerIleGlyValArgGluPheGluAspPheLeuLysIleHisLeuSer 290
QY 893 GAG-----GGAATATGCACA---CTGGAGACATGTAACACGCTTCAAGAAAGAGA 940
Db 291 GluThrCysAlaGluHisIleThrThrSerLeuSerAsnAspAspValMetLysGluAsn 310
QY 941 CTTGTGCCCAT--- 952
Db 311 LeuArgLysIleLeuAsnPheProLysAspAspLysLeuArgIleMetLeuGluGluAla 330
QY 952 ----- 952
Db 331 IleAspArgValLysLeuAsnThrArgArgLeuLeuArgGlnLysArgValSer 350
QY 953 ---GTCCCCCGCTGATGAGCTTACAGATCAGATC----- 988
Db 351 ArgLeuGluThrValPheGlyTyrPheAsnIleHisTyrIleAspAlaThrGluTyrIleLeu 370

```

```

OY 308 AATGAGCAACTGCTGATT----- 328
Db 82 GluValThrLeuAspLeuIleValLeuAsnSerProLeupheGluLeuAspProAsn 101
OY 329 -----GAGATATATTGCCCCGAGACAAATTCCTATTGTTGGGAGAAC 376
Db 102 PheHisAspHeuInLysIleArgAlaArgSerLysIleProValIleValGlyThr 121
OY 377 AATTATTACATGATCTGCTGTGAAA-----GTTCTGTCAATACCAACCCC 427
Db 122 ThrTyrTrpIleGluSerValLeuTyrGluAsnLeuIleGluThrThrSerAsp 141
OY 428 CAGAGAGTGGGCACTGGAAAGTATGACCGAAAGTGGAGCTGAAAAGAGATGCT 487
Db 142 AspValAspSerLysSerArgThrSerSerGluSerSerGluAspTrpGluGly 161
OY 488 CTT-----GTACTTCACAAACGCCCTAAGCCAGGTGGACCCAGAAATGGCTGCCAG 538
Db 162 IleSerAsnGlnGluLeuTrpAspGluLeuLysIleAspGluLysSerAlaLeuLeu 181
OY 539 CTCATCCACATGACAAACGAAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGA 598
Db 182 LeuHisProAsnAsnArgTyrArgValGlnArgAlaLeuGlnIlePheArgGluThrGly 201
OY 599 ATCTCATGTGATTTCTCCATCCGTCACATACGAAAGAGTGGTCCCTGCGA 658
Db 202 IleArgLysSerGluLeuValGluLysGlnLysSerAspLys---ThrValAspLeuGly 220
OY 659 GGTCTCTAGATTTCTCTAATCCCTTGACCTTGCTGCTATGACAGGAGCTTCTA 718
Db 221 GluArgLeuArgPheAspAsnSerLeuValIlePheMetAspAlaThrProGluValLeu 240
OY 719 GATGAGCGCTGGATTAAGAGGTTGATGACATCTGCTGCGCTCTTGAGGAACATA 778
Db 241 GlnGluArgLeuAspArgValAspLysMetIleLysLeuGlyLeuLysAsnGluLeu 260
OY 779 AGAGATTTCACAGAGCTTATTCAGAAAGATTTCCGAAATACCCAGACTATCAA 838
Db 261 IleGluPhe-----TyrAsnGlnHis-----AlaGluTyrIleAsnHisSerLys 275
OY 839 CATGATATCTTCAATCAATATGGCTTCAGAAATTCACAGTACCTG----- 886
Db 276 TyrGlyValMetGlnCysIleGlyLeuLysGlnPheValProTrpLeuAsnLeuAspPro 295
OY 887 -----ATCAGCTGAGGGAATGC-----ACA 907
Db 296 SerGluArgAspThrLeuAsnGlnLysAspLysLeuPheLysGlnLysCysAspValLys 315
OY 908 CTGGAGACT-----AGTAACGAGTTCTTAAG 934
Db 316 LeuHisThrArgGlnTyrAlaArgArgGlnArgGtrPyrArgSerArgLeuLeuLys 335
OY 935 AAGAGCACTGGTCCC---ATTGTCCCCCTGTATGCTTAAAGATGATGATGCTCG 991
Db 336 ArgSerAspGlnLysArgLysMetAlaSerThrLysMetLeuAspTrpSerAsp----- 353
OY 992 AATGGAGAGAGTCTGTTTGAACCTGCTTGAATCGTGAACATTCATCCAGGCG 1051
Db 354 -----LysTyrArgIleIleSerAspGlnMetAspIleValAspGlnTrpMetAsnGly 371
OY 1052 -----CAGAGGCTACAGGCACTCCAAATAAGTGCATACATGAAGTGAAGAC 1102
Db 372 IleAspLeuPheGlnAspIleSerThrAspThrAsn---ProIleLeuLysGlySerAsp 390
OY 1103 AAGAGAAGTATCACCTGTGTGACCTGTGATGCAATCATGATGGGAGTCCGATG 1162
Db 391 AlaAsnIleLeuLeuAsnGlnCysGlnIleCysAsnIleSerMetThrGlyLysAspAsnTrp 410
OY 1163 GAGGCGACATAAATCCAAATCCATGAAACCAATGAAAGAAAGAGA 1213
Db 411 GlnLysHisIleAspGlnLysLysHisLysHisIleAlaLysGlnLysLys 427

```

```

T38664
tRNA Isopentenyltransferase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
R:Murphy, L., Harris, D., Wood, V., Rajandram, M.A., Barrell, B.G.
Submitted to the EMBL Data Library, August 1997
A:Accession number: Z21804
A:Reference: T38664
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-434 <MUR>
A:Cross-references: EMBL:AL109739, NID:e1534774, PIDN:CA852278.1, GSPDB:GN00066, SPI
A:Experimental source: strain 972h-; cosmid C343
C:Genetics:
A:Gene: SPDB:SPAC343.15
A:Map position: 1

Alignment Scores:
Pred. No.: 8,46e-34 Length: 434
Score: 528.50 Matches: 141
Percent Similarity: 50.00% Conservative: 79
Best Local Similarity: 32.05% Mismatches: 133
Query Match: 14,788 Indels: 87
DB: 2 gaps: 19

US-09-513-151-3 (1-2041) x T38664 (1-434)
OY 80 CCTCTGTAGATTTCTGCGGGCCAGCGCAATCCAGCTGGGCTGACGTA 139
Db 4 ProLeuGlyValValIleGlyThrGlyAlaGlyLysSerAspLeuAlaValGlnLeu 23
OY 140 GCGCAGCGCTCGCGGTGAGATGTCAGCGCTGACTGCATGACAGTGTATGAAGCCTA 199
Db 24 AlaLysArgPheGlySerGlnValIleAsnAlaAspMetGlnIleTyrArgGlyPhe 43
OY 200 GACATCTACACAAAGAGTTTCTGCCAAGACAGAGAAATCTCCGGCACCACATGATC 259
Db 44 AspThrIleThrAsnLysIleThrValGlnGlnLysValHisIleArgLeuMet 63
OY 260 AGCTTTGTGAGCTCTTGTGACCAATTCACAGTGGGAGCTTCAGAAATGAGCAACT 319
Db 64 SerPheLeuAsn---PheAspLysGlnTyrSerValProGluPheGluArgAspIleSer 82
OY 320 GCTGTGATTAAGATATATTGCCCCGAGACAAATTCCTATTGTTGGAGGAACCAAT 379
Db 83 ArgValIleAspGlnIleHisSerGlnGlyLysIleProIleValIleGlyThrHis 102
OY 380 TATTCAATTGATCTGCTGTGAAA-----GTTCTGTGC 415
Db 103 TyrTrpLeuGlnSerLeuLeuPheGluAspThrThrLeuSerAlaIleAspLysLeuThr 122
OY 416 AAT-----ACCAAGCCCCAGAGATGGGCACTGGAAGATGATGACCGAAA 463
Db 123 AsnAspSerSerProSerLysProProHisProAspSerHis---IleLeuAsp----- 139
OY 464 GTGAGCTGAAAAGAGATGTCTTGTACTTCACAAAGCGCTTAAGCCAGTGGACCCA 523
Db 140 -----AspAspProSerAlaMetLeuSerTyrLeuLysLysIleAspPro 154
OY 524 GAATGGCTGCCAAGCTGCAATCCACATGACAAACGCAAAAGTGGCCAGAGCTTCAAGTT 583
Db 155 ValMetAlaGlnGlnTrpHisProArgAspThrArgLysIleArgAspSerLeuGlnIle 174
OY 584 TTGAGAAGAACAGAAATCTCATAGTGAATTCATGCTGTCACATGAGGAAGAGT 643
Db 175 TyrPheHisThrIleArgProProSerGlnIleTyrSerGlnGlnLysMetLysSerSer 194
OY 644 GGTGGTCCCTGGAGAGTCCCTGAAGTCTTAACCTTGCATCTTGGCTTCATGCT 703
Db 195 -----GlySerLysLeuArgTyrLysSer---LeuIlePheTrpAlaPheAla 209
OY 704 GACCAGGCACTTCTACATGAGCGCTTGATTAAGAGGTTGATGACATGCTTGGCTGGG 763

```

RESULT 3

## Alignment Scores:

Pred. No.: 2.38e-35 Length: 428  
 Score: 548.00 Matches: 143  
 Percent Similarity: 52.58% Conservative: 81  
 Best Local Similarity: 33.57% Mismatches: 144  
 Query Match: 15.33% Indels: 58  
 DB: 2 Gaps: 15

US-09-513-151-3 (1-2041) x S67176 (1-428)

```

QY 56 CTCAGGGC---CTGCAAGCGCCTACCTCTTGTAGTATCTCGGGCCAGGACACC 112
    ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Db 6 LeuysgIyCysLeuAsnMetSerIyLysValIleValIleAlGlyThrThrGlyVal 25
QY 113 GCGAATTCAGCGCTGGCTTGCAGCTAGGCGCGGCTGGGGTGAATCGTACGCT 172
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26 GlyLysSerGlnLeuSerIleGlnLeuValGlnLysPheAsnGlyGlnValIleAsnSer 45
QY 173 GACTCATGCAAGCTGTATGAAGGCTAGACATCATCACCAAGAGTTCTGCCAAGG 232
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 46 AspSerMetGlnValIyLysAspIleProIleIleThrAsnLysHisProLeuGlnGlu 65
QY 223 CAGAGAACTCGCCGCGCCACCATGATGAGCTTTGTGATCTCTTGACCAATTACACA 292
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 66 ArgGlnGlyIleProIleHisValMetAsnHisValAsp---TrpSerGlnGlyIyTrp 84
QY 293 GTGGGACTTCAGAAATAGACAACTGCTGATTAAGATATATTGGCCGAGACAAA 352
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 SerHisArgPheGlnIyThrGlyCysMetAsnAlaIleGlnAspIleHisArgGlyGly 104
QY 353 ATTCTATTGTTGGGAGGAACCAATATATCATTAATCTCTGCTCGGAAAGTCTT 412
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 IleProIleValIyGlyIyThrHisIyTrpIyLeuGlnIyThrPheAsnLys---Arg 123
QY 413 GTCAATACCAAGCCGAGAGAGATGGG---ACTGAGAAAGTATGATGACGGAAGAGTGAG 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 124 ValAspThrIyLysSerSerGlnArgIyLysLeuThrArgIyGlnLeuAsp-----Ile 140
QY 470 CTGAAAAGAGAGATGCTTGTACTTCACAAAGCCTAAGCAGGTGAGCCAGCAAAAG 529
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 LeuGlnSerThrAspProAspValIleIyThrAsnThrLeuValIyCysAspProAspIle 160
QY 530 GCTGCCAGCTGCATCCATCAGCAATAGCGCAAGTGGCCAGAGCTTGAAATTTTGA 589
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 AlaThrIyLysIyHisProAsnAspIyArgIyValGlnArgMetLeuGlnIyLeuIyTrp 180
QY 590 GAAACAGAAATCTCATATGATTTCTCCATCGTCAACATGAGGAAAGAGTGTGT 649
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 LysThrGlyLysIyLysProSerGlnIyThrPheAsnGlnGlnLysIle----- 195
QY 650 CCCCTTGAGAGCTCTGTGAAGTCTGTACACCTTGATCCCTTGGCTTCATGTCAGCAG 709
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 196 -----ThrLeuLysPhe---AspThrLeuPheLeuIyThrLeuIySerLysPro 210
QY 710 GCAGTTCTAGATGAGCGCTGTGATGAAGCGGTGATGACATCTTGCTGCTGCTTGG 769
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 211 GluProLeuPheGlnIyAspLeuAspAspIyValIyAspMetLeuGlnArgIyAlaLeu 230
QY 770 GAGGAAGTAAAGATTTTCCACAGAGCTATATATCAGAAAGTTCGAAATAGCCAG 829
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 231 GlnGlnIleLysGlnLeuIyGlnIyTrpIySerGlnAsnLysPheThr-----ProGlu 248
QY 830 GACTATCAACATGATATCTTCCATCAATATGGCTTCAAGAGAAATTCACAGACTGATC 889
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 249 GlnCysGlnAsnGlyValIyTrpGlnValIleGlyPheLysGlnPheLeuProIyLeu--- 267
QY 890 ACTGAGGAAAA-----TGC----- 904
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 ---ThrGlyLysThrAspAspAsnThrValLysLeuGlnAspCysIleGlnIyArgMetLys 286
QY 905 -----ACACTGGAGACTAGTACCAAGCTTCTAAAGAAAGACCTGTCCTCCATTCGCC 958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  
```

```

Db 287 ThrArgThrArgGlnIyThrAlaLysArgGlnValLysIyTrpIleLysLysMetLeuIlePro 306
QY 959 CCGTGTATGAGC-----TTAGAGTATGTGATGTCTCGAAGTGGAGAGAGCT 1006
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 AspIleLysGlyAspIleIyThrLeuLeuAspAlaThrAspLeuSerGlnIyTrpAspThrAsn 326
QY 1007 GTTCTGAAACCGCTCGTCAATTCGTGCAAAATTCATC-----CAG 1048
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 AlaSerGlnIyAlaIleIleAlaIleSerAsnAspIleIleSerAsnIyProIleLysGln 346
QY 1049 GCGCAACAGCCTACAGCCATCCAAATTAAGAGCCATCAATGACAGTGAACAAGAGA 1108
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 GluArgAlaProIyAlaLeuGlnIyLeuLeuSerIyGlyGlnIyThrMetLysLys 366
QY 1109 -----AGTTATCAGCTGTGTGACCTGTG-----GATCGA 1138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 LeuAspAspTrpThrHisIyTrpIyCysAsnValIyAspArgAsnAlaAspGlyLysAsnVal 386
QY 1139 ATCATCATTTGGGAGGAGCCGAAATGCGCAGCGCACATCAAAATCCCACTTGAACCA 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 ValAlaIleGlyGlnIyLysIyTrpLysIleHisLeuGlySerArgIyHisLysSerAsn 406
QY 1199 CTGAAGAAAAGAGAAGA 1216
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 LeuLysArgAsnThrArg 412

RESULT 2
T27538
hypochemical protein ZC395.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Connelly, M.
submitted to the EMBL Data Library, August 1994
A:Description: The sequence of C. elegans cosmid ZC395.
A:Accession: T27538
A:Reference number: Z20385
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-433 <CON>
A:Cross-references: EMBL:U13642; PIDN:AA853886.1; GSPDB:GN00021; CESP:ZC395.6
A:Experimental source: strain Bristol N2; clone ZC395
A:Genetic:
A:Gene: CESP:ZC395.6
A:Map position: 3
A:Introns: 43/1; 201/1; 266/3; 310/2; 341/3; 377/3; 410/2

Alignment Scores:
Pred. No.: 4.88e-34 Length: 433
Score: 531.50 Matches: 136
Percent Similarity: 50.34% Conservative: 84
Best Local Similarity: 31.12% Mismatches: 150
Query Match: 14.87% Indels: 67
DB: 2 Gaps: 14

US-09-513-151-3 (1-2041) x T27538 (1-433)
QY 71 CGGACCTTACTCTGTGTAGTATCTCGGGCCAGCGCAACGCAATCCAGCTGGCG 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 ArgThrAspProIleIlePheValIleGlyCysThrGlyThrGlyLysSerAspLeuGly 21
QY 131 TTGCAGCTAGGCGCAGCGCTCGCGGTGAGATCGTCACAGCGCTCATGTAGTCTAT 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 22 ValAlaIleAlaLysIyGlyIyGlnValIleIleSerValAspSerMetGlnIyThr 41
QY 191 GAAGCCTTACATATCATCAACAAGAGTTTGTGCCAAGACAGACAGATCTGCGCGCAG 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 LysGlyLeuAspIleAlaIyThrAsnLysIleThrGlnGlnIyGlnIyIleGlnHis 61
QY 251 CACAATGATCAGCTTGTGTGATCTCTT---GTGACCAATTAACAGAGTGTGATTCAGA 307
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 HisMetLysPheLeuAsnProSerGlnIySerSerIyAsnValHisSerPheArg 81
  
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 18:21:20 ; Search time 47.8883 Seconds  
(without alignments)  
8194.498 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 3575  
Sequence: 1 CTGCCATTAAGATGGCGTCG.....TTTACAGAAAAA 2041

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODL=frame+n2p model -DEV=xlh  
-Q=/cgn2\_1/USP10.spool/US09513151/unatc\_15042003\_141144\_26390/app-query.fasta\_1.2446  
-DB=PIR-73 -QEXT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=us09513151 @cgn.1.1.58 @runat.15042003.141144.26390 -NCPU=6 -ICPU=3  
-NO.XLPHY=NO.MMAP -LARGQUERY=NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: PIR.73:\*  
2: PIR3:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	548	15.3	428	2	S67176
2	531.5	14.9	433	2	T27538
3	528.5	14.8	434	2	T38664
4	493	13.8	466	2	T52061
5	471	13.2	461	2	P84676
6	417	11.7	314	2	P83945
7	403.5	11.3	314	2	G69657
8	397.5	11.1	305	2	AF1236
9	395.5	11.1	305	2	AC1599
10	378.5	10.6	309	2	D97126
11	348	9.7	305	2	C72366
12	344.5	9.6	339	2	A71473
13	339	9.5	311	2	B89905
14	338	9.5	318	2	T05569

15	334.5	9.4	311	2	F64046	delta(2)-isopenten
16	334	9.3	315	2	B82334	trna delta(2)-isop
17	333	9.3	294	2	B86703	trna isopentenyltr
18	329.5	9.2	325	2	AB2464	trna delta-2-isope
19	328.5	9.2	323	2	E83028	delta-2-isopenten
20	327.5	9.2	314	2	CG1735	trna delta-2-isope
21	324.5	9.1	314	2	S72942	trna isopentenyltr
22	324.5	9.1	311	2	F70505	probable trna delt
23	324	9.1	311	2	D97945	trna isopentenyltr
24	322.5	9.0	357	2	F66708	hypothetical prote
25	321.5	9.0	313	2	E81879	probable trna isop
26	320	9.0	294	2	G95077	trna isopentenylpy
27	319.5	8.9	342	2	E72019	trna delta-2-isope
28	319.5	8.9	342	2	AD0604	trna pyrophosphat
29	318.5	8.9	313	2	AD0046	trna isopentenyltr
30	314	8.8	305	2	G70391	trna delta-2-isope
31	313.5	8.8	330	2	T48100	trna isopentenyl t
32	312.5	8.7	317	2	H82848	trna delta(2)-isop
33	311	8.7	316	2	B37318	delta(2)-isopenten
34	310	8.7	316	2	C91272	hypothetical prote
35	310	8.7	316	2	C86113	hypothetical prote
36	305	8.5	312	2	T35111	probable trna delt
37	302	8.4	316	2	AH1048	trna isopentenyltr
38	301.5	8.4	306	2	D70202	2-methylthio-N6-is
39	286.5	8.0	306	2	H75366	trna delta-2-isope
40	281.5	7.9	363	2	H97781	hypothetical prote
41	280	7.8	303	2	S75554	trna delta-2-isope
42	278.5	7.8	313	2	AB1142	trna delta(2)-isop
43	275.5	7.7	326	2	G97604	hypothetical prote
44	272	7.6	298	2	B42643	trna isopentenyltr
45	272	7.6	298	2	A12826	trna delta(2)-isop

## ALIGNMENTS

## RESULT 1

S67176  
trna isopentenyltransferase (EC 2.5.1.8) - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein O5447w; protein YOR274w  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text change 21-Jul-2000  
C:Accession: S67176; A26717; S72045  
R:Cheret, G.; Sor, F.  
Submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67169  
A:Accession: S67176  
A:Molecule type: DNA  
A:Residues: 1-428 <CHP>  
A:Cross-references: EMBL:Z75182; NID:q1420613; PID:e252418; PID:q1420614; MIPS:YOR27  
A:Experimental source: strain S288C  
R:Naftarian, D.; Dihanich, M.E.; Martin, N.C.; Hopper, A.K.  
Mol. Cell. Biol. 7, 185-191, 1987  
A:Title: DNA sequence and transcript mapping of MOD5, features of the 5' region whic  
A:Reference number: A26717; MUID:87172703; PMID:3031457  
A:Accession: A26717  
A:Molecule type: DNA  
A:Residues: 1-374, 'R', 376-428 <RNAJ>  
A:Cross-references: EMBL:M15991  
R:Cheret, G.; Bernardi, A.; Sor, F.  
Yeast 12, 1059-1064, 1996  
A:Title: DNA sequence analysis of the VPB1-SNF2 region on chromosome XV of Saccharom  
A:Reference number: S72039; MUID:97051594; PMID:8896271  
A:Accession: S72045  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-428 <CHN>  
A:Cross-references: EMBL:X89633; NID:q1279694; PIDN:CA61780.1; PID:q1419759  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
C:Genetics:  
A:Gene: SGD:MOD5  
A:Cross-references: MIPS:YOR274w; SGD:S0005800  
A:Map position: 15R  
C:Keywords: transferase

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 18:52:25 ; Search time 63.5 Seconds  
(without alignments)

8565,812 Million cell updates/sec

Title: US-09-513-151-3

Sequence score: 1 CTGCCATAGATGGCTGCG.....TTTACAGAAAAA 2041

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1687582

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame+\_n2p.model -DEV-xlh  
-Q/cgn2.1/USPTO.spool/US09513151.runal.15042003.141246.27205/app.query.fasta.1.2183  
-DB-A.Geneseq.101002 -OFMT-fastan -SUFFIX-oln2p.rag -MINMATCH=0.1 -IOBCL=0  
-LOOEXT=0 -UNITS-bits -START=1 -END=1 MATRIX-oligo -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE-LOCAL  
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USBR=US09513151.ecgn.1.1.72-@runat.15042003.141246.27205 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAR -LARGEQUERY -NEG\_SCORES=0 -MAIT -LONGLOG -DEV.TIMEOUT=120  
-MAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A.Geneseq.101002.\*

1: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/genesep-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/genesep-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/genesep-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/genesep-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	38.4	411	23	ABB97295
2	123	18.6	222	22	ABB10278
3	123	18.6	222	22	ABB10476
4	123	18.6	222	22	AAU23398
5	48	7.3	57	22	ABB28879
6	48	7.3	57	22	ABB34055
7	48	7.3	57	22	AAE5837
8	48	7.3	57	22	AAE6220
9	48	7.3	57	22	AAE6220
10	48	7.3	57	22	AAE6220
11	16	2.4	221	20	AAU02534
12	11	1.4	94	22	AAU74967
13	9	1.2	11	19	AAU47911
14	8	1.2	11	19	AAU47917
15	8	1.2	11	19	AAU47901
16	8	1.2	12	19	AAU47905
17	8	1.2	48	22	AAE64616
18	8	1.2	60	22	AAE62297
19	8	1.2	101	23	ABP00506
20	8	1.2	118	22	AAU07994
21	8	1.2	120	22	AAU66805
22	8	1.2	132	21	AAU79156
23	8	1.2	136	23	ABP31089
24	8	1.2	212	22	AAU65505
25	8	1.2	216	20	AAU37775
26	8	1.2	219	22	AAU44657
27	8	1.2	234	22	AAU65782
28	8	1.2	269	22	AAU65782
29	8	1.2	277	22	AAU09572
30	8	1.2	330	23	ABP92775
31	8	1.2	330	23	ABP92775
32	8	1.2	351	22	ABP65281
33	8	1.2	351	22	ABP65281
34	8	1.2	357	23	ABP91502
35	8	1.2	378	21	AAU96592
36	8	1.2	378	21	AAU96592
37	8	1.2	378	22	AAU12381
38	8	1.2	378	22	AAU12381
39	8	1.2	378	22	AAU00666
40	8	1.2	378	22	AAU00666
41	8	1.2	378	22	AAU00666
42	8	1.2	378	22	AAU00666
43	8	1.2	378	22	AAU00666
44	8	1.2	378	22	AAU00666
45	8	1.2	378	23	AAU47914

## ALIGNMENTS

RESULT 1	ABB97295	standard; Protein: 411 AA.
ID	ABB97295	standard; Protein: 411 AA.
AC	ABB97295	standard; Protein: 411 AA.
DE	27-JUN-2002	(first entry)
DT	27-JUN-2002	(first entry)
XX	Novel human protein SEQ ID NO: 563.	
XX	Human; anti-inflamatory; anti-inflammatory; immunomodulator;	
XX	antifertility; cerebroprotective; cytoskeletal; gene therapy;	
XX	neuroprotective; antiparkinsonian; protein therapy; EST;	
XX	expressed sequence tag.	
XX	Homo sapiens.	
XX	OS	
XX	PN	WO200222660-A2.

PD 21-MAR-2002.  
XX 10-SEP-2001; 2001WO-US26015.  
XX 11-SEP-2000; 2000US-0659671.  
XX (HYSE-) HYSEQ INC.  
PA Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
XX WPI; 2002-292408/33.  
DR N-PSDB; ABN32481.  
XX  
PT An isolated polynucleotide for treating diseases associated with its  
PT encoded polypeptide such as cancer and multiple sclerosis.  
XX  
PS Example 2; SEQ ID NO 563; 509pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 444  
CC novel human proteins. These were isolated from expressed sequences tags  
CC (ESTs). They can be used to stimulate cell growth, to regulate  
CC hematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
CC Parkinson's disease. The present sequence is a protein of the invention.  
XX  
SQ Sequence 411 AA;  
  
Alignment Scores:  
Pred. No.: 4,69e-242 Length: 411  
Score: 254.00 Matches: 254  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.43% Indels: 0  
DB: 23 Gaps: 0  
  
US-09-513-151-3 (1-2041) x ABB97295 (1-411)  
QY 179 ATGCGAGCTGTGAAAGGCTGACATCATCACCAAGAGTTTCCCAAGACAGAGA 238  
DB 1 MetGlnValIYrGlGlyLeuAspIleIleThrAsnLysValSerAlaGlnGluInArg 20  
QY 239 ATCTGCCGACACACATGATGACCTTGTGATCCTCTTGACCAATTAACAGCTGTG 298  
DB 21 IleCysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyThrValVal 40  
QY 299 GACTTCAGAAATAGCAACTGCTGTGATTAATATTTGCCCGAGACAAATTCCT 358  
DB 41 AspPheArgAsnArgAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIlePro 60  
QY 359 ATTGTGGGGGAGGAACAATATTAATGATTCCTGCTGGAAGTTTCTTCAT 418  
DB 61 IleValValGlyGlyThrAsnTyTyIleGluSerLeuLeuTyPylsValLeuValAsn 80  
QY 419 ACCAAGCCCAAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGGAGCTTGAAG 478  
DB 81 ThrLysProGlnGluMetGlyThrGluLysValIleAspArgLysValGluLeuGluLys 100  
QY 479 GAGGATGCTTGTACTTCACAAAGCCTAAGCCAGGTGGACCCGAAATGGCTGCCAAG 538  
DB 101 GluAspArgLysLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaLys 120  
QY 539 CTGCATCCACATGACAAAGCAAGAGGGCCAGAGCTTGCAAGTTTGAAGAAACAGA 598  
DB 121 LeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluIleArgLys 140  
QY 599 ATCTGTACAGTAATTTCTCCATGCTGCAACATACGGAAGAGGTGGTCCCTTGA 658  
DB 141 IleSerHisSerGluPheLeuHisArgGlnHisThrGluGluGlyGlyGlyProLeuGly 160

QY 659 GGTCCTGAGAGTTCCTTAACCCCTTGATCCTTGCTCATGCTGACACGACAGTTCTA 718  
DB 161 GlyProLeuLysPheSerAsnProCysAlleuThrPheHisAlaAspGlnAlaValLeu 180  
QY 719 GATGAGCGCTTGATTAAGAGGGTGATGACATGCTCTGCTGGCTCTTGAGCAACTA 778  
DB 181 AspLysArgLeuAspLysArgValAspArgMetLeuAlaIleGlyLeuLeuGluGluLeu 200  
QY 779 AGAATTTTTCACACACCTTAATACAGAAATTTTCGGAAATATACCGAGACTATCAA 838  
DB 201 ArgAspPheHisArgArgTyLysGlnLysAsnValSerGluAsnSerGlnAspLysLeu 220  
QY 839 CATGTAATCTTCACATCAATATGCGCTTCAGAGAAATTTTCAGAGTACTGATCACTAGAGA 898  
DB 221 HisGlyIlePheGlnSerIleGlyPheLysGluPheHisGluThrGluGly 240  
QY 899 AATGACACTGTGAGACTAGTACCAAGCTTTAAAGAAAGCA 940  
DB 241 LysCysThrLeuGluThrSerAsnGlnLeuLeuLysGly 254  
  
RESULT 2  
ABBI0278  
ID ABB10278 standard; Protein: 222 AA.  
XX  
AC ABB10278;  
XX  
XX 10-JAN-2002 (first entry)  
XX  
DE Human CDNA SEQ ID NO: 586.  
XX  
XX Human; gene therapy; neural disorder; immune system disorder;  
KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
KW pulmonary disorder; cardiovascular disorder; renal disorder;  
KW proliferative disorder; Inflammation.  
XX Homo sapiens.  
XX WO200154474-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01349.  
XX  
XX 31-JAN-2000; 2000US-179065P.  
PR 04-FEB-2000; 2000US-180628P.  
PR 24-FEB-2000; 2000US-184664P.  
PR 02-MAR-2000; 2000US-186350P.  
PR 16-MAR-2000; 2000US-189874P.  
PR 17-MAR-2000; 2000US-190076P.  
PR 18-APR-2000; 2000US-198123P.  
PR 19-MAY-2000; 2000US-205515P.  
PR 07-JUN-2000; 2000US-209467P.  
PR 28-JUN-2000; 2000US-214886P.  
PR 30-JUN-2000; 2000US-215135P.  
PR 07-JUL-2000; 2000US-216647P.  
PR 07-JUL-2000; 2000US-216880P.  
PR 11-JUL-2000; 2000US-217487P.  
PR 11-JUL-2000; 2000US-217496P.  
PR 14-JUL-2000; 2000US-218290P.  
PR 26-JUL-2000; 2000US-220963P.  
PR 26-JUL-2000; 2000US-220964P.  
PR 14-AUG-2000; 2000US-224518P.  
PR 14-AUG-2000; 2000US-224519P.  
PR 14-AUG-2000; 2000US-225213P.  
PR 14-AUG-2000; 2000US-225214P.  
PR 14-AUG-2000; 2000US-225266P.  
PR 14-AUG-2000; 2000US-225267P.  
PR 14-AUG-2000; 2000US-225268P.  
PR 14-AUG-2000; 2000US-225270P.  
PR 14-AUG-2000; 2000US-225347P.  
PR 14-AUG-2000; 2000US-225757P.  
PR 14-AUG-2000; 2000US-225758P.

	PR	14-AUG-2000;	2000US-225759P.
	PR	18-AUG-2000;	2000US-226279P.
	PR	22-AUG-2000;	2000US-226681P.
	PR	22-AUG-2000;	2000US-226868P.
	PR	23-AUG-2000;	2000US-227182P.
	PR	30-AUG-2000;	2000US-227009P.
	PR	01-SEP-2000;	2000US-228924P.
	PR	01-SEP-2000;	2000US-229287P.
	PR	01-SEP-2000;	2000US-229343P.
	PR	01-SEP-2000;	2000US-229345P.
	PR	05-SEP-2000;	2000US-229509P.
	PR	05-SEP-2000;	2000US-229513P.
	PR	06-SEP-2000;	2000US-230437P.
	PR	06-SEP-2000;	2000US-230438P.
	PR	08-SEP-2000;	2000US-231242P.
	PR	08-SEP-2000;	2000US-231243P.
	PR	08-SEP-2000;	2000US-231244P.
	PR	08-SEP-2000;	2000US-231413P.
	PR	08-SEP-2000;	2000US-231414P.
	PR	08-SEP-2000;	2000US-232080P.
	PR	12-SEP-2000;	2000US-231968P.
	PR	14-SEP-2000;	2000US-232397P.
	PR	14-SEP-2000;	2000US-232398P.
	PR	14-SEP-2000;	2000US-232399P.
	PR	14-SEP-2000;	2000US-233400P.
	PR	14-SEP-2000;	2000US-233401P.
	PR	14-SEP-2000;	2000US-233063P.
	PR	14-SEP-2000;	2000US-233064P.
	PR	14-SEP-2000;	2000US-233065P.
	PR	21-SEP-2000;	2000US-234223P.
	PR	21-SEP-2000;	2000US-234274P.
	PR	25-SEP-2000;	2000US-234997P.
	PR	25-SEP-2000;	2000US-234998P.
	PR	26-SEP-2000;	2000US-235484P.
	PR	27-SEP-2000;	2000US-235834P.
	PR	27-SEP-2000;	2000US-235836P.
	PR	29-SEP-2000;	2000US-236327P.
	PR	29-SEP-2000;	2000US-236367P.
	PR	29-SEP-2000;	2000US-236368P.
	PR	29-SEP-2000;	2000US-236369P.
	PR	29-SEP-2000;	2000US-236370P.
	PR	02-OCT-2000;	2000US-237037P.
	PR	02-OCT-2000;	2000US-237038P.
	PR	02-OCT-2000;	2000US-237039P.
	PR	02-OCT-2000;	2000US-237040P.
	PR	13-OCT-2000;	2000US-239935P.
	PR	13-OCT-2000;	2000US-239937P.
	PR	20-OCT-2000;	2000US-240960P.
	PR	20-OCT-2000;	2000US-241221P.
	PR	20-OCT-2000;	2000US-241785P.
	PR	20-OCT-2000;	2000US-241786P.
	PR	20-OCT-2000;	2000US-241787P.
	PR	20-OCT-2000;	2000US-241808P.
	PR	20-OCT-2000;	2000US-241809P.
	PR	01-NOV-2000;	2000US-244617P.
	PR	08-NOV-2000;	2000US-246474P.
	PR	08-NOV-2000;	2000US-246475P.
	PR	08-NOV-2000;	2000US-246476P.
	PR	08-NOV-2000;	2000US-246477P.
	PR	08-NOV-2000;	2000US-246478P.
	PR	08-NOV-2000;	2000US-246523P.
	PR	08-NOV-2000;	2000US-246524P.
	PR	08-NOV-2000;	2000US-246525P.
	PR	08-NOV-2000;	2000US-246526P.
	PR	08-NOV-2000;	2000US-246527P.
	PR	08-NOV-2000;	2000US-246528P.
	PR	08-NOV-2000;	2000US-246532P.
	PR	08-NOV-2000;	2000US-246609P.
	PR	08-NOV-2000;	2000US-246610P.
	PR	14-AUG-2000;	2000US-225759P.
	PR	18-AUG-2000;	2000US-226279P.
	PR	22-AUG-2000;	2000US-226681P.
	PR	22-AUG-2000;	2000US-226868P.
	PR	23-AUG-2000;	2000US-227182P.
	PR	30-AUG-2000;	2000US-227009P.
	PR	01-SEP-2000;	2000US-228924P.
	PR	01-SEP-2000;	2000US-229287P.
	PR	01-SEP-2000;	2000US-229343P.
	PR	01-SEP-2000;	2000US-229345P.
	PR	05-SEP-2000;	2000US-229509P.
	PR	05-SEP-2000;	2000US-229513P.
	PR	06-SEP-2000;	2000US-230437P.
	PR	06-SEP-2000;	2000US-230438P.
	PR	08-SEP-2000;	2000US-231242P.
	PR	08-SEP-2000;	2000US-231243P.
	PR	08-SEP-2000;	2000US-231244P.
	PR	08-SEP-2000;	2000US-231413P.
	PR	08-SEP-2000;	2000US-231414P.
	PR	08-SEP-2000;	



QY	122	ACCTGGCGGTTTGACGTCAGGCGGCTGGGGGATGCGTCAGCGCTGACTCCATG	181
Db	41	ThrlenuAlaLeuGlnLeuGlnGlyGlnArgLeuGlyGlyGlnIleValSerAlaAspSerMet	60
QY	182	CAGTCATATAAAGCCCTAGACATCATCCACCAAGAAGTTTCTGCCCAAGAGAGAAATC	241
Db	61	GlnValIYrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle	80
QY	242	TGCCGCGACACATGATCAGCTTTGTTCGATCCTCTTGACCAATTCACAGTGGTGAC	301
Db	81	CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrrThrValAsp	100
QY	302	TTTCGAATAGACAACTGCTGATTCGATGAGATATATTTGCCGAGACAAATTCCTATT	361
Db	101	PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle	120
QY	362	GTTGTGGGA 370	
Db	121	ValValGly 123	
RESULT 4			
AAU23398	ID	AAU23398 standard; Protein; 222 AA.	
AAU23398;	AC		
18-DEC-2001 (first entry)	DT		
Novel human enzyme polypeptide #484.	DE		
Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;	KW		
lylase; hyperproliferative disorder; immunodeficiency disorder;	KM		
autoimmune disorder; neurological disorder; metabolic disorder;	KW		
inflammatory disorder; cardiovascular disorder; reproductive disorder;	KW		
blood-related disorder; infectious disorder; cytostatic; anti arthritic;	KW		
nephrotropic; anticoagulant.	KW		
Homo sapiens.	OS		
WO200155301-A2.	PN		
02-AUG-2001.	PD		
17-JAN-2001; 2001WO-US01239.	PF		
31-JAN-2000; 2000US-0179065.	PR		
04-FEB-2000; 2000US-0180628.	PR		
24-FEB-2000; 2000US-0184664.	PR		
02-MAR-2000; 2000US-0186350.	PR		
16-MAR-2000; 2000US-0189874.	PR		
17-MAR-2000; 2000US-0190076.	PR		
18-APR-2000; 2000US-0198123.	PR		
19-MAY-2000; 2000US-0205515.	PR		
07-JUN-2000; 2000US-0209467.	PR		
28-JUN-2000; 2000US-0214886.	PR		
30-JUN-2000; 2000US-0215135.	PR		
07-JUL-2000; 2000US-0216647.	PR		
07-JUL-2000; 2000US-0216880.	PR		
11-JUL-2000; 2000US-0217487.	PR		
11-JUL-2000; 2000US-0217486.	PR		
14-JUL-2000; 2000US-0218290.	PR		
26-JUL-2000; 2000US-0220963.	PR		
26-JUL-2000; 2000US-0220964.	PR		
14-AUG-2000; 2000US-0224518.	PR		
14-AUG-2000; 2000US-0224519.	PR		
14-AUG-2000; 2000US-0225213.	PR		
14-AUG-2000; 2000US-0225214.	PR		
14-AUG-2000; 2000US-0225267.	PR		
14-AUG-2000; 2000US-0225267.	PR		
14-AUG-2000; 2000US-0225268.	PR		
14-AUG-2000; 2000US-0225270.	PR		
14-AUG-2000; 2000US-0225447.	PR		



## Alignment Scores:

Pred. No.: 1.88e-112 Length: 222  
 Score: 123.00 Matches: 123  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 18.61% Indels: 0  
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x AAU23398 (1-222)

QY 2 TGCCATAAGATGGCGTCCGCGGCGTCCAGAGCAGTTCCTGTGGGAGTGCGCTCAG 61  
 Db 1 CysHISLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20  
 QY 62 GCGCTGCACGAGGACCTTACCTCTTGTAGTATTCGCGGCGCAGCGGACCGCAATCC 121  
 Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40  
 QY 122 ACGCTGGCGTTCAGAGCTAGGCGGCGGCGTCCGAGATGTCAGGCGTGCATCCATG 181  
 Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGluIleValSerAlaLysPserMet 60  
 QY 182 CAGGTCATAGAGGCTTACATCATCATCAACAAGTTTCCGCCAAGCAGCAGCAATC 241  
 Db 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80  
 QY 242 TGCCGCGCACCATGATCAGCTTGTGTGATCCCTTGTGACCAATTACACAGTGTGAC 301  
 Db 81 CysArgHISHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100  
 QY 302 TTCGAATATAGCAACATGCTGTGATTCAGATATATTTGCCCGAGCAAAATTCCTATT 361  
 Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaIleAspLysIleProIle 120  
 QY 362 GTGTGGGA 370  
 Db 121 ValValGly 123

RESULT 5  
 ABB28879  
 ID ABB28879 standard; Peptide: 57 AA.  
 AC ABB28879;  
 DT 01-FEB-2002 (first entry)  
 DE Peptide #1530 encoded by breast cell single exon nucleic acid probe.  
 KW Human; microarray; single exon probe; gene expression; breast;  
 KM disease; cancer.  
 OS Homo sapiens.  
 PN WO200157271-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001MO-US00662.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-46933/54.  
 XX

PT New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes  
 XX  
 XX Claim 27; SEQ ID NO 11847; 327Pp + sequence listing; English.  
 CC  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 4/4 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for  
 CC verifying the expression of regions of genomic DNA predicted to  
 CC encode proteins. They are useful for gene discovery, and for  
 CC determining predisposition and/or prognosing breast disease. Gene  
 CC expression analysis is useful for assessing the toxicity of chemical  
 CC agents on cells. The microarray of this invention presents a far greater  
 CC diversity of probes for measuring gene expression, with far less bias  
 CC than expressed sequence tag microarrays. The method is suitable for  
 CC rapid production of functional information from genomic sequence. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SO Sequence 57 AA;  
 Alignment Scores:  
 Pred. No.: 3.37e-38 Length: 57  
 Score: 48.00 Matches: 48  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.26% Indels: 0  
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x ABB28879 (1-57)  
 QY 185 GTCTATGAGGCTTACATCATCAACAAGTTTTCGCCAAGCAGCAAGATTCGC 244  
 Db 1 ValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIleGys 20  
 QY 245 CGGACACCATGATCAGCTTGTGTGATCCCTTGTGACCAATTACACAGTGTGACTTC 304  
 Db 21 ArgHISHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAspPhe 40  
 QY 305 AGAATAGCAACATGCTGTGATT 328  
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48

RESULT 6  
 ABB34055  
 ID ABB34055 standard; Peptide: 57 AA.  
 AC ABB34055;  
 DT 04-FEB-2002 (first entry)  
 DE Peptide #1561 encoded by human foetal liver single exon probe.  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.  
 KM Homo sapiens.  
 OS Homo sapiens.  
 PN WO200157277-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001MO-US00669.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 XX



PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human fetal liver -  
 PS Claim 27; SEQ ID NO 26690; 659bp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SQ Sequence 57 AA:  
  
 Alignment Scores:  
 Pred. No.: 3,37e-38 Length: 57  
 Score: 48.00 Matches: 48  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.26% Indels: 0  
 DB: 22 Gaps: 0  
  
 US-09-513-151-3 (1-2041) x ABB34055 (1-57)  
  
 QY 185 GTCRTATGAGGCTTGAGATATCATCCACAAAGGTTTCTGCCCAAGACAGAGATCTGC 244  
 DB 1 ValTYRGLUGLYLeuASPILeIleThrasnLysValSerAlaGlnGlnArgIleCys 20  
  
 QY 245 CGGCACCATGATGATGAGTTGTGGATCCTCTGTGACCAATTACACAGTGGAGACTTC 304  
 DB 21 ArgHSHSMetIleSerPheValAspProLeuValThrAsnTYThValValasphe 40  
  
 QY 305 AGAATAGAGCAACTGCTCTGATT 328  
 DB 41 ArgAsnArgAlaThrAlaLeuIle 48  
  
 RESULT 7  
 AAM54837  
 ID AAM54837 standard; Protein: 57 AA.  
 XX  
 AC AAM54837;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 26942.  
 XX  
 KW Human: brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 XX brains -  
 PS Example 4; SEQ ID NO: 26942; 650bp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 57 AA:  
  
 Alignment Scores:  
 Pred. No.: 3,37e-38 Length: 57  
 Score: 48.00 Matches: 48  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.26% Indels: 0  
 DB: 22 Gaps: 0  
  
 US-09-513-151-3 (1-2041) x AAM54837 (1-57)  
  
 QY 185 GTCRTATGAGGCTTGAGATATCATCCACAAAGGTTTCTGCCCAAGACAGAGATCTGC 244  
 DB 1 ValTYRGLUGLYLeuASPILeIleThrasnLysValSerAlaGlnGlnArgIleCys 20  
  
 QY 245 CGGCACCATGATGATGAGTTGTGGATCCTCTGTGACCAATTACACAGTGGAGACTTC 304  
 DB 21 ArgHSHSMetIleSerPheValAspProLeuValThrAsnTYThValValasphe 40  
  
 QY 305 AGAATAGAGCAACTGCTCTGATT 328  
 DB 41 ArgAsnArgAlaThrAlaLeuIle 48  
  
 RESULT 8  
 AAM67220  
 ID AAM67220 standard; Protein: 57 AA.  
 XX  
 AC AAM67220;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 27526.  
 XX  
 KW Human: bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0234687.  
 PR 04-OCT-2000; 2000US-0234687.  
 PR 04-OCT-2000; 2000US-0234687.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488901/53.  
 DR WPI; 2001-488901/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 PS Example 4; SEQ ID NO: 27526; 658bp + Sequence Listing; English.  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX Sequence 57 AA:  
 SQ  
 Alignment Scores:  
 Pred. No.: 3.37e-38 Length: 57  
 Score: 48.00 Matches: 48  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.26% Indels: 0  
 DB: Gaps: 0  
 US-09-513-151-3 (1-2041) x AAM67220 (1-57)  
 QY 185 GTCATGAGGCTAGACATCATCCACCAAGGTTTCTGCCAGACAGAGAAATCTGC 244  
 Db 1 ValTyrgluGlyLeuAspIleIleThrAsnLysValSerAlaGluGlnArgIleCys 20  
 QY 245 CGGCACACATGATCAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACTTC 304  
 Db 21 ArgHSHSMetIleSerPheValAspProLeuValThrAsnTyThrValValAspPhe 40  
 QY 305 AGAATAGAGCAACTGCTGTGATT 328  
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48  
 RESULT 9  
 AAM15066  
 ID AAM15066 standard; Protein; 57 AA.  
 AC AAM15066;  
 XX 12-OCT-2001 (first entry)  
 DE Peptide #1500 encoded by probe for measuring cervical gene expression.  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200157278-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US00670.  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0234687.  
 PR 04-OCT-2000; 2000US-0234687.  
 PR 04-OCT-2000; 2000US-0234687.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488901/53.  
 DR WPI; 2001-488901/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 PS Claim 27; SEQ ID NO 19892; 487bp; English.  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see A110068-A128459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 57 AA:  
 SQ  
 Alignment Scores:  
 Pred. No.: 3.37e-38 Length: 57  
 Score: 48.00 Matches: 48  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.26% Indels: 0  
 DB: Gaps: 0  
 US-09-513-151-3 (1-2041) x AAM15066 (1-57)  
 QY 185 GTCATGAGGCTAGACATCATCCACCAAGGTTTCTGCCAGACAGAGAAATCTGC 244  
 Db 1 ValTyrgluGlyLeuAspIleIleThrAsnLysValSerAlaGluGlnArgIleCys 20  
 QY 245 CGGCACACATGATCAGCTTTGTGATCCTCTTGACCAATTAACAGTGTGACTTC 304  
 Db 21 ArgHSHSMetIleSerPheValAspProLeuValThrAsnTyThrValValAspPhe 40  
 QY 305 AGAATAGAGCAACTGCTGTGATT 328  
 Db 41 ArgAsnArgAlaThrAlaLeuIle 48  
 RESULT 10  
 AAM27510  
 ID AAM27510 standard; Protein; 57 AA.  
 AC AAM27510;  
 XX 17-OCT-2001 (first entry)  
 DE Peptide #1547 encoded by probe for measuring placental gene expression.  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN WO200157272-A2.  
 PD 09-AUG-2001.  
 PF 30-JAN-2001; 2001WO-US00663.  
 PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 FI WPI; 2001-488897/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 PS  
 XX Claim 27; SEQ ID NO 27779; 654bp; English.  
 CC The present invention relates to single exon nucleic acid probes (SENPs;  
 CC see AAT31315-AAT57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 CC  
 SQ Sequence 57 AA:

Alignment Scores:  
 Pred. No.: 3.37e-38 Length: 57  
 Score: 48.00 Matches: 48  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.268 Indels: 0  
 DB: 22 Gaps: 0

US-09-513-151-3 (1-2041) x AAM27510 (1-57)

OY 185 GTCTATGAAGGCTGACATCATCAACAAAGTTTCTGCCCAAGACAGAGATCTGC 244  
 DB 1 VAIYTGUGLYLenuaspllelThrasnlysvaSerAlaGlnGlnArgrilecys 20  
 OY 245 CGGCACACATGATCAGCTTGTGGATCCTCTGTGACCAATTACAGAGTGGACTTC 304  
 DB 21 ARGHSHSMelileSerPheValaspProleuValThrAsnTyThValValasphe 40  
 OY 305 AGAATAGACCACTGCTCTGATT 328  
 DB 41 ArgAsnArgAlaThrAlaLeulle 48

RESULT 11

ID AAY02534 standard; Protein: 221 AA.  
 XX  
 AC AAY02534;  
 XX  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE Partial sequence of the Drosophila homologue of gro-1.  
 XX  
 KW gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;  
 KW hap-1 gene; cancer; aging; longevity; tumour formation;  
 KW physiological clock.  
 XX  
 OS Drosophila sp.  
 XX  
 PN WO9910482-A1.  
 XX  
 PD 04-MAR-1999.  
 XX  
 PF 20-AUG-1998; 98WO-CA00803.  
 XX  
 PR 25-AUG-1997; 97CA-2210251.  
 XX

PA (UYMC-) UNIV MCGILL.  
 XX  
 XX Barnes T, Hekim S, Lakowski B, Lemieux J;  
 FI WPI; 1999-190615/16.  
 DR  
 XX Molecular identity of the gro-1 gene - useful for cancer diagnosis  
 PT and/or prognosis, and where compounds affecting encoded proteins are  
 PT useful for enhancing longevity of a host and inhibiting tumour  
 PT formation  
 PS  
 XX Disclosure; Fig 10; 93pp; English.  
 XX

CC The specification describes the five genes of the *Caenorhabditis elegans*  
 CC gro-1 operon (AAX36071). The operon contains the gro-1 gene (AAX36072),  
 CC the gop-1 gene (AAX36074), the gop-2 gene (AAX36075), the gop-2 gene  
 CC (AAX36075), and the hap-1 gene (AAX36077). The gro-1 gene can be used in  
 CC a method for the diagnosis and/or prognosis of cancer in a patient.  
 CC Transgenic mice containing a gene knock-out of a murine gene homologue  
 CC of the gro-1 gene are useful as models of aging and cancer. The  
 CC proteins encoded by the genes are useful for identifying compounds  
 CC that affect the enzymatic activity of these proteins, in order to  
 CC enhance longevity of a host and inhibit tumour formation. The gro-1  
 CC gene, together with the gop-1, gop-2, gop-3 and hap-1 genes enables  
 CC study of a physiological clock. The present sequence represents a  
 CC partial sequence of the *Drosophila* homologue of gro-1.

SQ Sequence 221 AA:

Alignment Scores:  
 Pred. No.: 1.28e-06 Length: 221  
 Score: 16.00 Matches: 16  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.42% Indels: 0  
 DB: 20 Gaps: 0

US-09-513-151-3 (1-2041) x AAY02534 (1-221)

OY 356 CCTATTGTTGGAGGAGAACCAATTATTGATGATCTGCTGTG 403  
 DB 137 PROileValValGlyGlyThrasnTyrrileGlnSerLeuLeuTrp 152

RESULT 12

ID AAG74967 standard; Protein: 94 AA.  
 XX  
 AC AAG74967;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX  
 DE Human colon cancer antigen protein SEQ ID NO:5731.  
 XX  
 KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KW colorectal carcinoma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200122920-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 28-SEP-2000; 2000WO-US26524.  
 XX  
 PR 29-SEP-1999; 99US-0157137.  
 XX  
 PR 03-NOV-1999; 99US-0163280.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
 DR WPI; 2001-235357/24.  
 DR N-PSDB; AAH34372.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
PT useful for preventing, diagnosing and/or treating colorectal cancers -  
PS Claim 11; Page 7257; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing  
CC inactive proteins or to supplement the patient's own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.  
CC N. B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 94 AA:

Alignment Scores:  
Pred. No.: 11.9 Length: 94  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.36% Indels: 0  
DB: Gaps: 0

US-09-513-151-3 (1-2041) x AAG74967 (1-94)

QY 1017 CTGCTCTTGAATCGTCAAGTTCA 1043  
DB 11 LeuLeuLeuLysSerCysLysValSer 19

RESULT 13  
AAW47911  
ID AAW47911 standard; peptide: 11 AA.  
XX  
AC AAW47911;  
XX  
DT 09-JUN-1998 (first entry)  
XX  
DE Human tumour necrosis factor derived peptide 1168.  
XX  
KW Human: tumour necrosis factor; neutrophil; monocyte; macrophage;  
KW stimulatory activity; immunosuppressant; cancer; infection.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO9748725-A1.  
XX  
PD 24-DEC-1997.  
XX  
PF 20-JUN-1997; 97WO-AU00395.  
XX  
PR 29-OCT-1996; 96AU-0003309.  
PR 21-JUN-1996; 96AU-0000610.  
PR 06-SEP-1996; 96AU-0002165.  
XX  
XX (PEPT-) PEPT- LTD.  
XX  
XX Mack PC, Rathjen DA, Sleight JM, Widmer F.  
XX  
XX WPI; 1998-063077/06.

XX Tumour necrosis factor derived peptide(s) having neutrophil and/or  
PT monocyte-macrophage stimulatory activity - used for treating e.g.  
PT infections, immunosuppression or cancers  
XX  
PS Claim 4; Page 44; 65pp; English.

XX The present sequence represents a peptide derived from a human tumour  
CC necrosis factor peptide. The peptide can have neutrophil and/or  
CC monocyte/macrophage stimulatory activity. The peptide is derived from  
CC the sequence of tumour necrosis factor (TNF)-derived peptide 419  
CC (PETHYLRHRI; see AU74762/91 and 44664/93). It can have improved  
CC properties such as increased potency, extended in vivo half life or,  
CC particularly, specificity of action. It can have neutrophil stimulatory  
CC activity (class 1), equal neutrophil and monocyte/macrophage stimulatory  
CC activity (class 2), or preferentially enhanced monocyte/macrophage  
CC stimulatory activity (class 3). It can be used for the treatment or  
CC prevention of infections. In particular, it can be used for treating  
CC AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic  
CC fibrosis, community acquired pneumonia, meningitis, Mycobacteria,  
CC Chlamydia, Brucellae, Francisella, Pasteurella, Legionellosis,  
CC Histoplasmosis, Listeriosis, Pneumocystis carinii, Trypanosoma cruzi,  
CC coccidian parasitical infection, an inherited primary neutropenic  
CC disorder, an inherited defect of phagocytic cell function, an acquired  
CC defect of phagocytic cell function, immunosuppression due to the  
CC administration of immunosuppressive drugs, and other bacterial, fungal,  
CC viral or protozoan infection, infectious mononucleosis, paroxysmal  
CC nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft  
CC versus host disease.

XX SQ Sequence 11 AA:

Alignment Scores:  
Pred. No.: 150 Length: 11  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.22% Indels: 0  
DB: Gaps: 0

US-09-513-151-3 (1-2041) x AAW47911 (1-11)

QY 1967 ACAGCAGTGATPACTACACACA 1944  
DB 3 ThrAlaValIleIleThrHisPThr 10

RESULT 14  
AAW47917  
ID AAW47917 standard; peptide: 11 AA.  
XX  
AC AAW47917;  
XX  
DT 09-JUN-1998 (first entry)  
XX  
DE Human tumour necrosis factor derived peptide 1175.  
XX  
KW Human: tumour necrosis factor; neutrophil; monocyte; macrophage;  
KW stimulatory activity; immunosuppressant; cancer; infection.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO9748725-A1.  
XX  
PD 24-DEC-1997.  
XX  
PF 20-JUN-1997; 97WO-AU00395.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 1  
XX /note="nipoctic acid"

PR 29-OCT-1996; 96AU-0003309.  
PR 21-JUN-1996; 96AU-0000610.  
PR 06-SEP-1996; 96AU-0002165.  
XX  
XX (PEPT-) PEPTech LTD.  
XX  
PI Mack PO, Rathjen DA, Sleigh JM, Widmer F;  
XX  
XX WPI: 1998-063077/06.  
XX  
XX Tumour necrosis factor derived peptide(s) having neutrophil and/or  
PT monocytic-macrophage stimulatory activity - used for treating e.g.  
PT infections, immunosuppression or cancers  
XX  
XX  
PS Claim 4; Page 44; 65pp; English.  
XX  
XX The present sequence represents a peptide derived from a human tumour  
CC necrosis factor peptide. The peptide can have neutrophil and/or  
CC monocytic-macrophage stimulatory activity. The peptide is derived from  
CC the sequence of tumour necrosis factor (TNF)-derived peptide 419  
CC (PSTHVLITHTI; see AU74762/91 and 44664/93). It can have improved  
CC properties such as increased potency, extended in vivo half life or,  
CC particularly, specificity of action. It can have neutrophil stimulatory  
CC activity (class 1), equal neutrophil and monocytic-macrophage stimulatory  
CC activity (class 2), or preferentially enhanced monocytic-macrophage  
CC stimulatory activity (class 3). It can be used for the treatment or  
CC prevention of infections. In particular, it can be used for treating  
CC AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic  
CC fibrosis, community acquired pneumonia, meningitis, Mycobacteriella,  
CC Chlamydia, Brucellae, Francisella, Pasteurellosis, Legionellosis,  
CC Histoplasmosis, listeriosis, Pneumocystis carni, Trypanosoma cruzi,  
CC coccidian parasitical infection, an inherited primary neutropenic  
CC disorder, an inherited primary defect of phagocytic cell function, an  
CC inherited secondary defect of phagocytic cell function, an acquired  
CC defect of phagocytic cell function, immunosuppression due to the  
CC administration of immunosuppressive drugs, and other bacterial, fungal,  
CC viral or protozoan infection, infectious mononucleosis, paroxysmal  
CC nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft  
CC versus host disease.  
XX  
XX  
SQ Sequence 11 AA:  
  
Alignment Scores:  
Pred. No.: 150 Length: 11  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.22% Indels: 0  
DB: 19 Gaps: 0  
  
US-09-513-151-3 (1-2041) x AAM47917 (1-11)  
OY 1967 ACAGCAGTGATATACTACACACA 1944  
DB 3 ThrAlaValIleIleThrHisThr 10  
|||||  
RESULT 15  
AAM47901  
ID AAM47901 standard; peptide; 11 AA.  
XX  
XX AAM47901;  
XX  
XX 09-JUN-1998 (first entry)  
XX  
XX Human tumour necrosis factor derived peptide 1100.  
DE  
XX  
XX Human tumour necrosis factor; neutrophil; monocytic; macrophage;  
KW stimulatory activity; immunosuppressant; cancer; infection.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX WO9748725-A1.  
PN

XX  
XX 24-DEC-1997.  
PD  
XX  
XX 20-JUN-1997; 97WO-AU00395.  
PF  
XX  
XX 29-OCT-1996; 96AU-0003309.  
PR 21-JUN-1996; 96AU-0000610.  
PR 06-SEP-1996; 96AU-0002165.  
XX  
XX (PEPT-) PEPTech LTD.  
XX  
PI Mack PO, Rathjen DA, Sleigh JM, Widmer F;  
XX  
XX WPI: 1998-063077/06.  
XX  
XX Tumour necrosis factor derived peptide(s) having neutrophil and/or  
PT monocytic-macrophage stimulatory activity - used for treating e.g.  
PT infections, immunosuppression or cancers  
XX  
XX  
PS Claim 4; Page 44; 65pp; English.  
XX  
XX The present sequence represents a peptide derived from a human tumour  
CC necrosis factor peptide. The peptide can have neutrophil and/or  
CC monocytic-macrophage stimulatory activity. The peptide is derived from  
CC the sequence of tumour necrosis factor (TNF)-derived peptide 419  
CC (PSTHVLITHTI; see AU74762/91 and 44664/93). It can have improved  
CC properties such as increased potency, extended in vivo half life or,  
CC particularly, specificity of action. It can have neutrophil stimulatory  
CC activity (class 1), equal neutrophil and monocytic-macrophage stimulatory  
CC activity (class 2), or preferentially enhanced monocytic-macrophage  
CC stimulatory activity (class 3). It can be used for the treatment or  
CC prevention of infections. In particular, it can be used for treating  
CC AIDS, cancer, diabetes, nosocomial infection, tuberculosis, cystic  
CC fibrosis, community acquired pneumonia, meningitis, Mycobacteriella,  
CC Chlamydia, Brucellae, Francisella, Pasteurellosis, Legionellosis,  
CC Histoplasmosis, listeriosis, Pneumocystis carni, Trypanosoma cruzi,  
CC coccidian parasitical infection, an inherited primary neutropenic  
CC disorder, an inherited primary defect of phagocytic cell function, an  
CC inherited secondary defect of phagocytic cell function, an acquired  
CC defect of phagocytic cell function, immunosuppression due to the  
CC administration of immunosuppressive drugs, and other bacterial, fungal,  
CC viral or protozoan infection, infectious mononucleosis, paroxysmal  
CC nocturnal, haemoglobinuria, leukaemia, lymphoma, myelofibrosis, or graft  
CC versus host disease.  
XX  
XX  
SQ Sequence 11 AA:  
  
Alignment Scores:  
Pred. No.: 150 Length: 11  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.22% Indels: 0  
DB: 19 Gaps: 0  
  
US-09-513-151-3 (1-2041) x AAM47901 (1-11)  
OY 1967 ACAGCAGTGATATACTACACACA 1944  
DB 3 ThrAlaValIleIleThrHisThr 10  
|||||  
Search completed: April 21, 2003, 19:14:30  
Job time : 66.5 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:51:16 ; Search time 13.7893 Seconds

(without alignments)  
8416.087 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 164

Sequence: 1 TGTGACCTCTGTGTCGAAAT.....TGAACCAACTGAGAGAAAAGA 90

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 9138288

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:  
-O=/cgn2\_1/USPRO.spool/US09513151.runtat\_15042003\_141145\_26426/app\_query.fasta\_1.2446  
-DB=Pending\_Patents\_AA\_Main.qcmt-fastan -SUFFIX=n2p.ram -MIMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62  
-TRANS=human40.cdt -LIST=45 -DCCALIGN=200 -THR\_SCORE=pet -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09513151.ecgn.1.1.422.ecgnat.15042003\_141145\_26426  
-NCU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG\_SCORES=0 -WATT -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -RGAPOP=6  
-RGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending\_Patents\_AA\_Main:\*  
1: /cgn2\_6/prodata/1/paa/PCRTUS.COMB.pep:\*  
2: /cgn2\_6/prodata/1/paa/US06.COMB.pep:\*  
3: /cgn2\_6/prodata/1/paa/US07.COMB.pep:\*  
4: /cgn2\_6/prodata/1/paa/US080.COMB.pep:\*  
5: /cgn2\_6/prodata/1/paa/US081.COMB.pep:\*  
6: /cgn2\_6/prodata/1/paa/US082.COMB.pep:\*  
7: /cgn2\_6/prodata/1/paa/US083.COMB.pep:\*  
8: /cgn2\_6/prodata/1/paa/US084.COMB.pep:\*  
9: /cgn2\_6/prodata/1/paa/US085.COMB.pep:\*  
10: /cgn2\_6/prodata/1/paa/US086.COMB.pep:\*  
11: /cgn2\_6/prodata/1/paa/US087.COMB.pep:\*  
12: /cgn2\_6/prodata/1/paa/US088.COMB.pep:\*  
13: /cgn2\_6/prodata/1/paa/US089.COMB.pep:\*  
14: /cgn2\_6/prodata/1/paa/US090.COMB.pep:\*  
15: /cgn2\_6/prodata/1/paa/US091.COMB.pep:\*  
16: /cgn2\_6/prodata/1/paa/US092.COMB.pep:\*  
17: /cgn2\_6/prodata/1/paa/US093.COMB.pep:\*  
18: /cgn2\_6/prodata/1/paa/US094.COMB.pep:\*  
19: /cgn2\_6/prodata/1/paa/US095.COMB.pep:\*  
20: /cgn2\_6/prodata/1/paa/US096.COMB.pep:\*  
21: /cgn2\_6/prodata/1/paa/US097.COMB.pep:\*  
22: /cgn2\_6/prodata/1/paa/US098.COMB.pep:\*  
23: /cgn2\_6/prodata/1/paa/US099.COMB.pep:\*  
24: /cgn2\_6/prodata/1/paa/US100.COMB.pep:\*  
25: /cgn2\_6/prodata/1/paa/US101.COMB.pep:\*  
26: /cgn2\_6/prodata/1/paa/US102.COMB.pep:\*  
27: /cgn2\_6/prodata/1/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	143	21 US-09-757-028-2199	Sequence 2199, Ap
2	164	100.0	143	26 US-10-222-911-2199	Sequence 2199, Ap
3	164	100.0	465	1 PCT-US02-07826-151	Sequence 151, Ap
4	164	100.0	465	24 US-10-097-340-151	Sequence 151, Ap
5	164	100.0	467	1 PCT-US02-07826-153	Sequence 153, Ap
6	164	100.0	467	24 US-10-097-340-153	Sequence 153, Ap
7	64	39.0	294	27 US-60-389-987-773	Sequence 773, Ap
8	64	39.0	294	27 US-60-412-418-773	Sequence 773, Ap
9	64	39.0	635	27 US-60-167-217-10879	Sequence 10879, A
10	64	39.0	884	27 US-60-173-464-8824	Sequence 8824, Ap
11	63	38.4	617	1 PCT-US01-18569-3108	Sequence 3108, Ap
12	63	38.4	617	26 US-10-264-049-3108	Sequence 3108, Ap
13	63	38.4	845	26 US-10-205-219-133	Sequence 133, App
14	63	38.4	847	23 US-09-976-594-401	Sequence 401, App
15	63	38.4	847	25 US-10-153-668-450	Sequence 450, App
16	63	38.4	847	27 US-60-240-409-401	Sequence 401, App
17	63	38.4	864	1 PCT-US01-08631-49060	Sequence 49060, A
18	61.5	37.5	354	25 US-10-179-131-10074	Sequence 10074, A
19	61.5	37.5	386	16 US-09-248-796-17815	Sequence 17815, A
20	61.5	37.5	386	27 US-60-096-409-17815	Sequence 17815, A
21	61	37.2	314	20 US-09-614-150-24477	Sequence 24477, A
22	61	37.2	314	27 US-60-191-637-24572	Sequence 24572, A
23	61	37.2	314	27 US-60-191-681-19349	Sequence 19349, A
24	60	36.6	251	26 US-10-219-999-57395	Sequence 57395, A
25	60	36.6	289	21 US-09-708-427-56054	Sequence 56054, A
26	60	36.6	324	21 US-09-708-427-56053	Sequence 56053, A
27	60	36.6	430	19 US-09-513-151-2	Sequence 2, App11
28	60	36.6	466	21 US-09-708-427-56052	Sequence 56052, A
29	59	36.0	428	21 US-09-791-537-46482	Sequence 46482, A
30	58	35.8	72	19 US-09-513-996A-73087	Sequence 73087, A
31	58	35.8	72	19 US-09-513-996A-78196	Sequence 78196, A
32	58	35.8	121	19 US-09-513-996A-73086	Sequence 73086, A
33	58	35.8	121	19 US-09-513-996A-78195	Sequence 78195, A
34	58	35.4	578	20 US-09-614-150-1971	Sequence 1971, Ap
35	58	35.4	578	27 US-60-191-637-1977	Sequence 1977, Ap
36	58	35.4	578	27 US-60-191-681-1578	Sequence 1578, Ap
37	58	35.4	884	20 US-09-614-150-10866	Sequence 10866, A
38	58	35.4	884	27 US-60-191-637-10898	Sequence 10898, A
39	58	35.4	884	27 US-60-191-681-8546	Sequence 8546, Ap
40	57.5	35.1	935	20 US-09-614-150-1641	Sequence 1641, Ap
41	57.5	35.1	935	27 US-60-173-464-22515	Sequence 22515, A
42	57.5	35.1	935	27 US-60-191-637-1647	Sequence 1647, Ap
43	57.5	35.1	935	27 US-60-191-681-1332	Sequence 1332, Ap
44	57.5	35.1	956	27 US-60-161-932-1331	Sequence 1331, Ap
45	57	35.2	126	21 US-09-758-472-4966	Sequence 4966, Ap

#### ALIGNMENTS

RESULT 1  
US-09-757-028-2199  
Sequence 2199, Application US/09757028  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PM001  
CURRENT FILING DATE: 2001-01-09  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
NUMBER OF SEQ ID NOS: 2660  
SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 2199
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-757-028-2199 (1-143)

Alignment Scores:
Pred. No.: 1,97e-15      Length: 143
Score: 164.00           Matches: 30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 21                  Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-757-028-2199 (1-143)

QY 1 TGTGACCTCTGTGATCGATCATGATGGGATCGCGAATGGCGACGCACATAAATCC 60
    |||||||
Db 73 Cysaspleucysaspargillellellyasprgltutrpalaiahlsilleysser 92

QY 61 AAATCCCACTGGAACCACTGAAGAAAGA 90
    |||||||
Db 93 LysSerHisLeuAsnGlnLeuLysLysArg 102

RESULT 2
US-10-222-911-2199
; Sequence 2199, Application US/10222911
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM001C1N
; CURRENT APPLICATION NUMBER: US/10/222,911
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/757,028
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 2660
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2199
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-911-2199

Alignment Scores:
Pred. No.: 1,97e-15      Length: 143
Score: 164.00           Matches: 30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 26                  Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-222-911-2199 (1-143)

QY 1 TGTGACCTCTGTGATCGATCATGATGGGATCGCGAATGGCGACGCACATAAATCC 60
    |||||||
Db 73 Cysaspleucysaspargillellellyasprgltutrpalaiahlsilleysser 92

QY 61 AAATCCCACTGGAACCACTGAAGAAAGA 90
    |||||||
Db 93 LysSerHisLeuAsnGlnLeuLysLysArg 102

RESULT 3
PCT-US02-07826-151
; Sequence 151, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
```

```
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07826-151

Alignment Scores:
Pred. No.: 2,17e-15      Length: 465
Score: 164.00           Matches: 30
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 1                    Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US02-07826-151 (1-465)

QY 1 TGTGACCTCTGTGATCGATCATGATGGGATCGCGAATGGCGACGCACATAAATCC 60
    |||||||
Db 395 Cysaspleucysaspargillellellyasprgltutrpalaiahlsilleysser 414

QY 61 AAATCCCACTGGAACCACTGAAGAAAGA 90
    |||||||
Db 415 LysSerHisLeuAsnGlnLeuLysLysArg 424

RESULT 4
US-10-097-340-151
; Sequence 151, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangt KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
```

```

; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-151

Alignment Scores:
Pred. No.: 2,17e-15      Length: 465
Score: 164.00           Matches: 30
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 24                  Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-097-340-151 (1-465)
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGAGGCGACACATAAAATCC 60
Db 395 Cysaspleucysaspargillellellellysparglutrpalalaahisileysse 414
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
Db 415 LysSerHisLeuAnGlnLeuLysArg 424

RESULT 5
PCT-US02-07826-153
; Sequence 153, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07826-153

Alignment Scores:
Pred. No.: 2,17e-15      Length: 467
Score: 164.00           Matches: 30
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 1                  Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-097-340-153 (1-467)
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGAGGCGACACATAAAATCC 60
Db 397 Cysaspleucysaspargillellellellysparglutrpalalaahisileysse 416
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
Db 417 LysSerHisLeuAnGlnLeuLysArg 426

RESULT 6
US-10-097-340-153
; Sequence 153, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCHE
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATIS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEV
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, JR.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xunlei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-153

Alignment Scores:
Pred. No.: 2,17e-15      Length: 467
Score: 164.00           Matches: 30
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 24                  Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-097-340-153 (1-467)
OY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGGAGGCGACACATAAAATCC 60
Db 397 Cysaspleucysaspargillellellellysparglutrpalalaahisileysse 416
OY 61 AATCCCACTTGAAACCACTGAAGAAAGA 90
```





```
Percent Similarity: 57.14% Conservative: 5
Best Local Similarity: 39.29% Mismatches: 12
Query Match: 27 Indels: 0
DB: 27 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-173-464-8824 (1-884)

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGCGACATTAATCC 60
    ||| :|||||
Db 255 CysG1bLencysAspValThrCysThrGlyThrAspAlaGlyAlaHisValArgGly 274
    ||| :|||||

QY 61 AAATCCACCTTGAACCACTGAG 84
    ||| ||| :|||||
Db 275 AlalysHisGlnAsnValValys 282
    ||| ||| :|||||

RESULT 11
PCT-US01-18569-3108
; Sequence 3108, Application PC/TUS0118569
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA133PCr
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,467
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3108
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (64)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-18569-3108

Alignment Scores:
Pred. No.: 4.08 Length: 617
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US01-18569-3108 (1-617)

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGCGACATTAATCC 60
    ||| :|||||
Db 69 CysSerIleCysAspLeuProValHisSerAsnIysGlnThrSerGlnHisIleAsnGly 88
    ||| :|||||

QY 61 AAATCCAC 69
    |||||
Db 89 AlaserHis 91
    |||||

RESULT 12
US-10-264-049-3108
; Sequence 3108, Application US/10264049
; GENERAL INFORMATION:
; APPLICANT: Birtse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 3108
; LENGTH: 617

TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (64)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3108

Alignment Scores:
Pred. No.: 4.08 Length: 617
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 26 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-264-049-3108 (1-617)

QY 1 TGTGACCTCTGTGATCGAATCATCATTTGGGATCGCGAATGGCGACGCGACATTAATCC 60
    ||| :|||||
Db 293 CysSerIleCysAspLeuProValHisSerAsnIysGlnThrSerGlnHisIleAsnGly 312
    ||| :|||||

QY 61 AAATCCAC 69
    |||||
Db 313 AlaserHis 315
    |||||

RESULT 14
US-09-594-401
; Sequence 401, Application US/09976594
; GENERAL INFORMATION:
```

```

APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
PRIOR FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 401
LENGTH: 847
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No: 3676741CD1
US-09-976-594-401

```

```

Alignment Scores:
Pred. No.: 4.19 Length: 847
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 23 Gaps: 0

```

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-09-976-594-401 (1-847)

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGCGCAGCCACATATAATCC 60

DB 293 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleasncl 312

OY 61 AAATCCAC 69

DB 313 AlaserHis 315

RESULT 15

US-10-153-668-450

Sequence 450, Application US/10153668

GENERAL INFORMATION:

APPLICANT: HONDA, Goichi

APPLICANT: MATSUDA, Akio

APPLICANT: MURAMATSU, Shuji

APPLICANT: ISHIZAWA, Kenya

TITLE OF INVENTION: State Activating Gene

FILE REFERENCE: 1254-0207P

CURRENT APPLICATION NUMBER: US/10/153,668

CURRENT FILING DATE: 2002-05-24

PRIOR APPLICATION NUMBER: US 60/293,172

PRIOR FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: US 60/316,031

PRIOR FILING DATE: 2001-08-31

PRIOR APPLICATION NUMBER: US 60/328,403

PRIOR FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: JP 2001-157043

PRIOR FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: JP 2001-260681

PRIOR FILING DATE: 2001-08-30

PRIOR APPLICATION NUMBER: JP 2001-313175

PRIOR FILING DATE: 2001-10-10

NUMBER OF SEQ ID NOS: 488

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 450

LENGTH: 847

TYPE: PRT

ORGANISM: Homo sapiens

US-10-153-668-450

Alignment Scores:

Pred. No.: 4.19

Score: 63.00

Percent Similarity: 60.87%

Length: 847

Matches: 9

Conservative: 5

```

Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 25 Gaps: 0

```

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-10-153-668-450 (1-847)

OY 1 TGTGACCTCTGTGATCGAATCATTCATTTGGGATCGCGAATGCGCAGCCACATATAATCC 60

DB 293 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleasncl 312

OY 61 AAATCCAC 69

DB 313 AlaserHis 315

Search completed: April 21, 2003, 19:12:19  
Job time: 15.7893 secs



```
Db 35 LysSerHisLeuAsnGlnLeuLysLysArg 44

RESULT 2
US-60-453-050-7969
; Sequence 7969, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7969
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-7969

Alignment Scores:
Pred. No.: 1.66e-15 Length: 85
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-050-7969 (1-85)

QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCGACCATATAAATCC 60
Db 15 CysAspLeuCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 34

QY 61 AAATCCCACTTGAACCACTGACAGCAAGAAAGA 90
Db 35 LysSerHisLeuAsnGlnLeuLysLysArg 44

RESULT 3
US-60-453-135-7967
; Sequence 7967, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7967
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-7967

Alignment Scores:
Pred. No.: 1.73e-15 Length: 221
Score: 164.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-135-7967 (1-221)

QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCGACCATATAAATCC 60
Db 151 CysAspLeuCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 170

QY 61 AAATCCCACTTGAACCACTGACAGCAAGAAAGA 90

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-380-731-563 (1-411)

QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCGAATGGCGACCATATAAATCC 60
Db 341 CysAspLeuCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 360
```



```
Db 95 LyslyshsAlaAsnLysValIysArg 103
RESULT 9
US-10-218-140-3448
; Sequence 3448, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218, 140
; PRIORITY FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curanator Version 1.0
; SEQ ID NO 3448
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-140-3448

Alignment Scores:
Pred. No.: 0.967 Length: 302
Score: 64.00 Matches: 10
Percent Similarity: 65.52% Conservative: 9
Best Local Similarity: 34.48% Mismatches: 10
Query Match: 39.02% Indels: 0
Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-218-140-3448 (1-302)
QY 1 TGTGACCTCTGTGATCGATCATTCATTTGGGATCGCGCATATAAATCC 60
Db 83 CyslysvAlcYcysAlaLeuLeuIleSerGlnLysLeuAlaHisTyrGlnSer 102
QY 61 AAATCCCACTTGAACCACTGAAGAAA 87
Db 103 LyslyshsAlaAsnLysValIysArg 111

RESULT 10
US-60-452-680-16040
; Sequence 16040, Application US/60452680
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: GRUPE, Andrew
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01450
; CURRENT APPLICATION NUMBER: US/60/452,680
; PRIORITY FILING DATE: 2003-03-07
; NUMBER OF SEQ ID NOS: 116213
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16040
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(294)
; OTHER INFORMATION: Xaa - Any Amino Acid
US-60-452-680-16040

Alignment Scores:
Pred. No.: 1.36 Length: 294
```

```
Score: 63.00 Matches: 8
Percent Similarity: 60.87% Conservative: 6
Best Local Similarity: 34.78% Mismatches: 9
Query Match: 38.41% Indels: 0
Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-452-680-16040 (1-294)
QY 1 TGTGACCTCTGTGATCGATCATTCATTTGGGATCGCGCATATAAATCC 60
Db 231 CysserilcysAspLysValPheAspLeuLysAspTyrGlnLysValIysGly 250
QY 61 AAATCCCACTTGAACCACTGAAGAAA 87
Db 251 LysleuHis 253

RESULT 11
PCT-US03-07002-35
; Sequence 35, Application PC/TUS0307002
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; RICHARDSON, Thomas W.;
; APPLICANT: ELLIOTT, Vicki S.; SPRAGUE, William W.;
; APPLICANT: JIANG, Xin; TANG, Y. Tom;
; APPLICANT: ZEBARJADIAN, Yeganeh; BAUGHN, Mariah R.;
; APPLICANT: JACKSON, Alan A.; KABLE, Amy E.;
; APPLICANT: LEE, Ernestine A.; KHAN, Farrah A.;
; APPLICANT: YDE, Henry; FORSYTHE, Ian J.;
; APPLICANT: RAMKUMAR, Jayalaxmi; GRIFFIN, Jennifer A.;
; APPLICANT: LI, Joana X.; MARQUIS, Joseph P.;
; APPLICANT: LEHR-MASON, Patricia M.; JIN, Pei;
; APPLICANT: HARKINS, Phillips R.; WILSON, Amy D.;
; APPLICANT: SWARNAKAR, Anita; HE, Ann;
; APPLICANT: HAFALIA, April J.A.; TRAN, Bao;
; APPLICANT: DUGGAN, Brendan M.; EMERLING, Brooke M.;
; APPLICANT: BOROMSKY, Mark L.; YAO, Monique G.;
; APPLICANT: CHAWLA, Narinder K.; BURFORD, Neil;
; APPLICANT: KHARE, Reena; LEE, Sally;
; APPLICANT: BECHA, Shanya D.; LEE, Soo Yeun
; TITLE OF INVENTION: NUCLEIC ACID-ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1377 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/07002
; PRIORITY FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/362,329
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/364,438
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/373,891
; PRIOR FILING DATE: 2002-04-19
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 7509991CD1
PCT-US03-07002-35

Alignment Scores:
Pred. No.: 1.36 Length: 339
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x PCT-US03-07002-35 (1-339)
QY 1 TGTGACCTCTGTGATCGATCATTCATTTGGGATCGCGCATATAAATCC 60
Db 293 CysserilcysAspLeuPheValHisSerAsnLysGlnTyrPserGlnHisIleAsnGly 312
```

```
QY 61 AATCCAC 69
Db 313 AlaserHis 315

RESULT 12
US-60-453-135-8780
; Sequence 8780, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8780
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8780

Alignment Scores:
Pred. No.: 142 Length: 847
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
Db: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-135-8780 (1-847)
QY 1 TGTGACCTCTGTGATCATCATTCATGCGGATCGGATCGGACGACATTAATCC 60
||| ::::| ||| ::::| ||| ::::| ||| ::::| ||| ::::| |||
Db 293 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleAsnGly 312
QY 61 AATCCAC 69
Db 313 AlaserHis 315

RESULT 13
US-60-453-050-8780
; Sequence 8780, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8780
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8780

Alignment Scores:
Pred. No.: 142 Length: 847
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
Db: 7 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-60-453-050-8780 (1-847)
QY 1 TGTGACCTCTGTGATCATCATTCATGCGGATCGGATCGGACGACATTAATCC 60
||| ::::| ||| ::::| ||| ::::| ||| ::::| ||| ::::| |||
Db 293 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleAsnGly 312
QY 61 AATCCAC 69
Db 313 AlaserHis 315

RESULT 14
US-09-949-016-11423
; Sequence 11423, Application US/09949016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11423
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11423

Alignment Scores:
Pred. No.: 142 Length: 853
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best Local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
Db: 5 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-949-016-11423 (1-853)
QY 1 TGTGACCTCTGTGATCATCATTCATGCGGATCGGATCGGACGACATTAATCC 60
||| ::::| ||| ::::| ||| ::::| ||| ::::| ||| ::::| |||
Db 299 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerGlnHisIleAsnGly 318
QY 61 AATCCAC 69
Db 319 AlaserHis 321

RESULT 15
US-09-675-784A-9300
; Sequence 9300, Application US/09675784A
; GENERAL INFORMATION:
; APPLICANT: HARE, ROBERTA S.
; APPLICANT: SHAW, KAREN J.
; APPLICANT: SHIMER JR., GEORGE H.
; APPLICANT: KESSLER, MARCO
; APPLICANT: NOLLING, JORR
; APPLICANT: ZENG, QIANDONG
; TITLE OF INVENTION: ASPERGILLUS FUMIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 2976-4020US1
; CURRENT APPLICATION NUMBER: US/09/675,784A
; CURRENT FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/156,338
; PRIOR FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 13925
; SEQ ID NO 9300
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-675-784A-9300

Alignment Scores:
```



Pred. No.:	10.3	length:	242
Score:	57.00	Matches:	7
Percent Similarity:	53.57%	Conservative:	8
Best Local Similarity:	25.00%	Mismatches:	13
Query Match:	34.76%	Indels:	0
DB:	5	Gaps:	0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) X US-09-675-784A-9300 (1-242)

[illegible]

Search completed: April 21, 2003, 18:58:58  
Job time : 4.07555 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:49:25 ; Search time 2.57626 Seconds

(without alignments)  
5283.055 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGACCACTGAAGAAGAAGA 90

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 288829 seqs, 75613885 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+.n2p.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09513151/runat\_15042003\_141145\_26460/app\_query.fasta\_1.2446  
-DB=Published Applications\_AA -GWT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -SMART=1 -END=1 -MATRIX=blcosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=pco -NORM=ext -HEA5SIDE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09513151\_@CGN\_1.1.15\_runat\_15042003\_141145\_26460  
-NCPU=6 -ICPU=3 -NO\_XLPEX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	57	34.8	737	9	US-10-102-806-731
2	54	32.9	123	9	US-09-986-480-246
3	54	32.9	187	9	US-09-986-480-368
4	53.5	32.6	238	10	US-09-764-864-901

#### ALIGNMENTS

RESULT 1  
US-10-102-806-731  
Sequence 731, Application US/10102806  
Publication No. US20030054421A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA103PIC1  
CURRENT APPLICATION NUMBER: US/10/102,806  
CURRENT FILING DATE: 2002-03-22  
PRIOR APPLICATION NUMBER: 09/925,298  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05881  
PRIOR FILING DATE: 2000-03-08  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 846  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 731  
LENGTH: 737  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-102-806-731  
Alignment Scores:  
Pred. No.: 3.53  
Score: 57.00  
Percent Similarity: 53.57%  
Length: 737  
Matches: 11  
Conservative: 4

```

Best Local Similarity: 39.29%      Mismatches: 13
Query Match: 34.76%      Indels: 0
DB: 9      Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-10-102-806-731 (1-737)

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCCGCGATGGCGCCACATTAATCC 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 CysGluLeuCySAspCysValValAlaAspAlaTyrAlaAlaHisIleArgGly 66
OY 61 AATGCCACTTGAACCACTGAAG 84
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 AlalysHisGlnLysValValLys 74

RESULT 2
US-09-986-480-246
; Sequence 246, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 246
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-246

Alignment Scores:
Pred. No.: 7.92      Length: 123
Score: 54.00      Matches: 12
Percent Similarity: 43.18%      Conservative: 7
Best Local Similarity: 27.27%      Mismatches: 11
Query Match: 32.93%      Indels: 14
DB: 9      Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-986-480-246 (1-123)

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCCGCGATGGCGCCACATTAATCC 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 CysAsnValCySAspCysValValLysAspSerIleAsnPhelLeuAspHisIleAsnGly 25
OY 61 AATGCCAC-----TTGAACCA 78
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 LysLysHisGlnArgAsnLeuGlyMetSerMetArgValGluArgSerThrLeuAspGln 45
OY 79 CTGAAGAAAGA 90
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46 ValLysLysArg 49

RESULT 3
US-09-986-480-368
; Sequence 368, Application US/09986480
; Publication No. US20030027999A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: PS500P1
; CURRENT APPLICATION NUMBER: US/09/986,480
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 368
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-986-480-368

Alignment Scores:
Pred. No.: 8.38      Length: 187
Score: 54.00      Matches: 12
Percent Similarity: 43.18%      Conservative: 7
Best Local Similarity: 27.27%      Mismatches: 11
Query Match: 32.93%      Indels: 14
DB: 9      Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-986-480-368 (1-187)

OY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCCGCGATGGCGCCACATTAATCC 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 70 CysAsnValCySAspCysValValLysAspSerIleAsnPhelLeuAspHisIleAsnGly 89
OY 61 AATGCCAC-----TTGAACCA 78
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 90 LysLysHisGlnArgAsnLeuGlyMetSerMetArgValGluArgSerThrLeuAspGln 109
OY 79 CTGAAGAAAGA 90
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 ValLysLysArg 113

RESULT 4
US-09-764-864-901
; Sequence 901, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 901
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (153)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (185)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (195)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (196)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (210)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-901

Alignment Scores:
Pred. No.: 10.3      Length: 238
Score: 53.50      Matches: 13

```

;  
PRIOR FILING DATE: 2001-08-10

US-09-925-302-679 (1-179) x US-09-925-302-679 (1-90) COPY\_1121\_1210

TYPE: PRT

ORGANISM: Homo sapiens

US-09-999-248-2

## Alignment Scores:

Pred. No.:	47.4	Length:	578
Score:	49.50	Matches:	8
Percent Similarity:	50.00%	Conservative:	5
Best Local Similarity:	30.77%	Mismatches:	6
Query Match:	30.56%	Indels:	7
DB:	9	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-09-999-248-2 (1-578)

QY 89 CTTTCTTCATGTCG-----TTCAAGTGGATTGCAATTGAT 51

DB 457 LeuYrPheIIleTyrPAlCysArGAspIleGInserPheArGTrPheAlAspLeu 476

QY 50 TGGCGTGGCCCATTCGCGA 33

DB 477 CysMetLeuHIsAsnLys 482

## RESULT 12

US-10-101-464A-114

; Sequence 114, Application US/10101464A

; Publication No. US20030046728A1

## GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signalling

; FILE REFERENCE: 11000.1020c2

; CURRENT FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: US/10101464A

; PRIOR FILING DATE: 2002-03-18

; PRIOR APPLICATION NUMBER: 09/704,302

; PRIOR FILING DATE: 2000-11-01

; PRIOR APPLICATION NUMBER: 09/228,986

; PRIOR FILING DATE: 1999-01-12

; PRIOR APPLICATION NUMBER: 60/162,866

; PRIOR FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US00/00724

; PRIOR FILING DATE: 2000-01-11

; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 114

; LENGTH: 762

; TYPE: PRT

; ORGANISM: Pampus radiata

US-10-101-464A-114

## Alignment Scores:

Pred. No.:	69.9	Length:	762
Score:	48.50	Matches:	10
Percent Similarity:	63.16%	Conservative:	2
Best Local Similarity:	52.63%	Mismatches:	6
Query Match:	29.57%	Indels:	1
DB:	9	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-10-101-464A-114 (1-762)

QY 15 TCGAATCATTCATTCGGGATCGCGATGCGACATTAATCCAAATCCACTT 71

DB 261 SerAsnHisHisTrp---SerLeuAlaGlyProValAlaAlaIleArGValProLeu 278

## RESULT 13

US-10-105-695-2

; Sequence 2, Application US/10105695

; Publication No. US20020197272A1

## GENERAL INFORMATION:

; APPLICANT: Galloway, Darrel R.

; APPLICANT: Mateczun, Alfred J.

; TITLE OF INVENTION: Methods for Protection Against Lethal Infection with Bacillus Ant

; FILE REFERENCE: 22727/04115

; CURRENT APPLICATION NUMBER: US/10105,695

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER: US 09/747,521

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 809

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-10-105-695-2

## Alignment Scores:

Pred. No.:	70.4	Length:	809
Score:	48.50	Matches:	9
Percent Similarity:	53.85%	Conservative:	5
Best Local Similarity:	34.62%	Mismatches:	7
Query Match:	29.57%	Indels:	5
DB:	9	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-10-105-695-2 (1-809)

QY 24 CATTGGGATTCGCGAATG-----GGCAGCGCACATTAATCCAAATCCCA 68

DB 313 HisTrpSerAspSerLeuSerGluGluGlyArgGlyLeuLeuLysLysLeuGlnIlePro 332

QY 69 CTGGAACCACTGAGAA 86

DB 333 IleGluProLysLysAsp 338

## RESULT 14

US-10-105-694-2

; Sequence 2, Application US/10105694

; Publication No. US20030003109A1

## GENERAL INFORMATION:

; APPLICANT: Mateczun, Alfred J.

; APPLICANT: Galloway, Darrel R.

; TITLE OF INVENTION: Methods for Protection Against Lethal Infection with Bacill

; FILE REFERENCE: 22727/04116

; CURRENT APPLICATION NUMBER: US/10105,694

; CURRENT FILING DATE: 2002-03-25

; PRIOR APPLICATION NUMBER: US 09/747,521

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 809

; TYPE: PRT

; ORGANISM: Bacillus anthracis

US-10-105-694-2

## Alignment Scores:

Pred. No.:	70.4	Length:	809
Score:	48.50	Matches:	9
Percent Similarity:	53.85%	Conservative:	5
Best Local Similarity:	34.62%	Mismatches:	7
Query Match:	29.57%	Indels:	5
DB:	9	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-10-105-694-2 (1-809)

QY 24 CATTGGGATTCGCGAATG-----GGCAGCGCACATTAATCCAAATCCCA 68

DB 313 HisTrpSerAspSerLeuSerGluGluGlyArgGlyLeuLeuLysLysLeuGlnIlePro 332

QY 69 CTGGAACCACTGAGAA 86

DB 333 IleGluProLysLysAsp 338

## RESULT 15

US-09-747-521-2

; Sequence 2, Application US/09747521

; Patent No. US20020051791A1

; GENERAL INFORMATION:

APPLICANT: Galloway, Darrel  
APPLICANT: Mateczun, Alfred  
TITLE OF INVENTION: Methods for Protection Against Lethal Infection with Bacillus Anthracis  
FILE REFERENCE: 22727/04079  
CURRENT APPLICATION NUMBER: US/09/747,521  
CURRENT FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO: 2  
LENGTH: 809  
TYPE: PRT  
ORGANISM: Bacillus anthracis  
US-09-747-521-2

## Alignment Scores:

Pred. No.:	70.4	length:	809
Score:	48.50	Matches:	9
Percent Similarity:	53.85%	Conservative:	5
Best Local Similarity:	34.62%	Mismatches:	7
Query Match:	29.57%	Indels:	5
DB:	10	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-09-747-521-2 (1-809)

QY	24	CATTGGGATCGCGAATG-----GGCAGCGCACATAAATCCAATCCA	68
Db	313	HistpSerAspserLeuSerGluGluGlyArgGlyLeuLeuLysLysLeuGlnIlePro	332
QY	69	CTTGACCACTGACAGA	86
Db	333	IleGluProLysLysAsp	338

Search completed: April 21, 2003, 19:01:37  
Job time : 2.57626 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:41:35 ; Search time 0.971375 Seconds  
(without alignments)  
5452.195 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 164  
Sequence: 1 TGTGACCTCTGTGATGCAAT.....TGACCACTGAGAGAAAAGA 90

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlh  
-O=/cgn2\_1/USPRO.spool/US09513151/runat\_15042003\_141144\_26406/app.query.fasta\_1.2446  
-DB=Issued\_Patents\_AA -OPMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151\_ECGN\_1\_1\_23\_runat\_15042003\_141144\_26406 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfill1.pcp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	32.9	500	US-09-141-000-2	Sequence 2, Appl1
2	53	32.3	418	US-09-141-000-6	Sequence 6, Appl1
3	53	32.3	431	US-08-836-620A-19	Sequence 19, Appl1
4	53	32.3	435	US-09-040-508-2	Sequence 2, Appl1
5	53	32.3	435	US-09-500-654-2	Sequence 2, Appl1
6	53	32.3	438	US-09-141-000-4	Sequence 4, Appl1
7	51	31.1	653	US-09-443-184-50	Sequence 50, Appl1
8	49	30.2	136	US-09-288-143-190	Sequence 190, Appl1
9	49	30.2	226	US-08-679-493A-23	Sequence 23, Appl1
10	48.5	29.6	762	US-09-228-986-114	Sequence 114, Appl1
11	48.5	29.6	776	US-08-021-601-2	Sequence 2, Appl1
12	48.5	29.6	776	US-08-082-849B-2	Sequence 2, Appl1

13	48.5	29.6	776	5	PCT-US94-01624-2	Sequence 2, Appl1
14	48	29.6	878	4	US-09-735-934A-2	Sequence 2, Appl1
15	47.5	29.0	244	3	US-09-135-782-4	Sequence 4, Appl1
16	47.5	29.0	244	4	US-09-193-191-4	Sequence 4, Appl1
17	46.5	28.4	220	2	US-08-726-306A-29	Sequence 29, Appl1
18	46.5	28.4	317	1	US-07-709-949-2	Sequence 2, Appl1
19	46	28.0	31	4	US-08-602-999A-44	Sequence 44, Appl1
20	46	28.0	31	4	US-08-278-865-44	Sequence 44, Appl1
21	46	28.0	31	4	US-09-500-114-44	Sequence 38, Appl1
22	46	28.4	170	4	US-08-444-818-38	Sequence 4, Appl1
23	46	28.0	216	2	US-08-737-716-4	Sequence 2, Appl1
24	46	28.4	382	4	US-08-969-815-2	Sequence 2, Appl1
25	46	28.4	382	4	US-09-120-025-2	Sequence 2, Appl1
26	46	28.4	382	4	US-09-710-481-2	Sequence 2, Appl1
27	46	28.0	418	4	US-09-202-893B-4	Sequence 4, Appl1
28	46	28.0	437	2	US-08-737-716-2	Sequence 4, Appl1
29	46	28.0	502	4	US-09-330-970-1	Sequence 1, Appl1
30	46	28.4	516	4	US-08-867-611-6	Sequence 6, Appl1
31	46	28.4	516	5	PCT-US92-06965A-11	Sequence 11, Appl1
32	46	28.0	518	2	US-08-836-620A-18	Sequence 18, Appl1
33	46	28.0	568	1	US-08-262-424-7	Sequence 7, Appl1
34	46	28.0	568	2	US-08-493-187-7	Sequence 7, Appl1
35	46	28.0	568	2	US-08-717-587A-4	Sequence 4, Appl1
36	46	28.0	568	3	US-08-883-610A-4	Sequence 4, Appl1
37	46	28.0	568	4	US-08-936-094A-4	Sequence 4, Appl1
38	46	28.0	568	5	PCT-US95-07844-7	Sequence 7, Appl1
39	46	28.4	798	4	US-08-867-611-36	Sequence 36, Appl1
40	46	28.4	894	4	US-09-735-934A-4	Sequence 4, Appl1
41	46	28.4	1011	4	US-08-850-328-5	Sequence 5, Appl1
42	46	28.4	1011	4	US-08-444-818-54	Sequence 54, Appl1
43	46	28.4	2261	4	US-08-444-818-66	Sequence 66, Appl1
44	46	28.4	2436	4	US-08-444-818-75	Sequence 75, Appl1
45	46	28.4	2772	4	US-08-444-818-89	Sequence 89, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-141-000-2  
Sequence 2, Application US/09141000  
Patent No. 6054295  
GENERAL INFORMATION:  
APPLICANT: Chen, Pang  
TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR  
FILE REFERENCE: 19999Y  
CURRENT APPLICATION NUMBER: US/09/141,000  
CURRENT FILING DATE: 1998-08-26  
NUMBER OF SEQ ID NOS: 30  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 500  
TYPE: PRT  
ORGANISM: Human  
US-09-141-000-2

Alignment Scores:  
Pred. No.: 2.4  
Score: 54.00  
Percent Similarity: 57.14%  
Best Local Similarity: 50.00%  
Query Match: 32.93%  
DB: 3  
Gaps: 2

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-09-141-000-2 (1-500)

QY 1 TGTGACCTCTGTGATGCA-----ATCACCATTGGGAGATCCGATGCGCAGCCAGC 51  
|||||  
DB 244 CysAspLeuAlaAspArgGluLeuValAlaIleIleGly-----TrrAlaYshIs 260  
|||||  
QY 52 ATTAATCAATCCCACTGAGC 75  
|||||  
DB 261 IleProGlyPheSerSerLeuSer 268  
|||||



```
RESULT 2
US-09-141-000-6
; Sequence 6, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chem, Pang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 199999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-6

Alignment Scores:
Pred. No.: 3.34 Length: 418
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.008 Mismatches: 6
Query Match: 32.328 Indels: 6
DB: Gaps: 2

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-141-000-6 (1-418)

QY 1 TGTGACCTCTGTGATCGA-----ATCATCATTTGGGAGATCGCAATGGCAGCGCAC 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 269 CysAspLeuAlaAspArgGluLeuValIleIleGly-----TrpAlaLysHis 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 52 ATAAATCCAAATCCCACTTGAAAC 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 IleProGlyPheSerThrLeuSer 293

RESULT 3
US-08-836-620A-19
; Sequence 19, Application US/08836620A
; Patent No. 5958710
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Orphan receptor
; NUMBER OF SEQUENCES: 19
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,620A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/03933
; FILING DATE:
; APPLICATION NUMBER: GB 9518272.1
; FILING DATE: 08-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9605550.4
; FILING DATE: 15-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9607532.0
; FILING DATE: 11-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9609576.5
; FILING DATE: 08-MAY-1996
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-836-620A-19

Alignment Scores:
Pred. No.: 3.36 Length: 431
Score: 53.00 Matches: 11
Percent Similarity: 58.33% Conservative: 3
Best Local Similarity: 45.83% Mismatches: 10
Query Match: 32.328 Indels: 0
DB: Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-08-836-620A-19 (1-431)

QY 1 TGTGACCTCTGTGATCGATCATCATTTGGGATCGCAATGGCAGCGCAATAATCC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 244 CysAspLeuAlaAspArgGluLeuValIleIleSerTrpAlaLysHisIleProGly 263
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 AAATCCCACTTG 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 264 PheSerAsnLeu 267

RESULT 4
US-09-040-508-2
; Sequence 2, Application US/09040508
; Patent No. 6069239
; GENERAL INFORMATION:
; APPLICANT: Mathias, Stephen L.
; TITLE OF INVENTION: NOVEL COMPOUNDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,508
; FILING DATE: 17-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9705451.4
; FILING DATE: 17-MAR-1997
; APPLICATION NUMBER: GB 9803289.9
; FILING DATE: 16-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-30008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-040-508-2

Alignment Scores:
Pred. No.: 3.37 Length: 435
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.008 Mismatches: 6
```

```

Query Match: 32.32% Indels: 6
DB: 3 Gaps: 2

US-09-513-151-3.COPY_1121_1210 (1-90) x US-09-040-508-2 (1-435)
QY 1 TGTGACCTGTGTATCGA-----ATCATCATTTGGGATGCGGAATGGCGACGAC 51
   |||||  |||||  ::|||  |||||  |||||  |||
Db 246 CysaspLeuAlaaspArgGluLeuValIleIleGly-----TrrAlaIysHis 262
QY 52 ATAAATCCAAATCCCACTTGAAAC 75
   |||  |||  |||  |||  |||
Db 263 IleProGlyPheSerThrLeuSer 270

RESULT 5
US-09-500-654-2
; Sequence 2, Application US/09500654
; Patent No. 6359116
; GENERAL INFORMATION:
; APPLICANT: Stephen L. Mathias
; TITLE OF INVENTION: No. 6359116el Compounds
; FILE REFERENCE: GP-30008-D1
; CURRENT APPLICATION NUMBER: US/09/500,654
; CURRENT FILING DATE: 2000-02-09
; EARLIER APPLICATION NUMBER: 9705451.4
; EARLIER FILING DATE: 1997-03-17
; EARLIER APPLICATION NUMBER: 9803289.9
; EARLIER FILING DATE: 1998-02-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-500-654-2

Alignment Scores:
Pred. No.: 3.37 Length: 435
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 32.32% Indels: 6
DB: 4 Gaps: 2

US-09-513-151-3.COPY_1121_1210 (1-90) x US-09-500-654-2 (1-435)
QY 1 TGTGACCTGTGTATCGA-----ATCATCATTTGGGATGCGGAATGGCGACGAC 51
   |||||  |||||  ::|||  |||||  |||||  |||
Db 246 CysaspLeuAlaaspArgGluLeuValIleIleGly-----TrrAlaIysHis 262
QY 52 ATAAATCCAAATCCCACTTGAAAC 75
   |||  |||  |||  |||  |||
Db 263 IleProGlyPheSerThrLeuSer 270

RESULT 6
US-09-141-000-4
; Sequence 4, Application US/09141000
; Patent No. 6054295
; GENERAL INFORMATION:
; APPLICANT: Chen, Pang
; TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
; FILE REFERENCE: 19999Y
; CURRENT APPLICATION NUMBER: US/09/141,000
; CURRENT FILING DATE: 1998-08-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Human
US-09-141-000-4

Alignment Scores:

```

```

Pred. No.: 3 4 Length: 458
Score: 53.00 Matches: 14
Percent Similarity: 57.14% Conservative: 2
Best Local Similarity: 50.00% Mismatches: 6
Query Match: 32.32% Indels: 6
DB: 3 Gaps: 2

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-141-000-4 (1-458)
QY 1 TGTGACCTCTGTGATCGA-----ATCATCATTTGGGGATGCGAATGGCGAGCGAC 51
||||| ||||| :||| ||||| ||||| |||
Db 269 CysaspLeuAlaaspArgGluLeuValValIleleely-----TrpAlaLysHis 285
||| ||| ||| ||| |||
QY 52 ATAAATCCAAATCCCACTTGAC 75
||| ||| ||| ||| |||
Db 286 IleProGlyPheSerThrLeuSer 293

RESULT 7
US-09-443-184-50
; Sequence 50, Application US/09443184A
; Patent No. 6372431
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Zweiger, Gary
; APPLICANT: Kaser, Matthew R.
; APPLICANT: Panzer, Scott
; APPLICANT: Selhammer, Jeffrey J.
; APPLICANT: Yue, Henry
; APPLICANT: Baughn, Mariah
; APPLICANT: Azimzal, Yalda
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
; FILE REFERENCE: PC-0007 US
; CURRENT APPLICATION NUMBER: US/09/443,184A
; CURRENT FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PERL Program
; SEQ ID NO 50
; LENGTH: 653
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6372431 3511087CD1
US-09-443-184-50

Alignment Scores:
Pred. No.: 7 57 Length: 653
Score: 51.00 Matches: 7
Percent Similarity: 56.52% Conservative: 6
Best Local Similarity: 30.43% Mismatches: 10
Query Match: 31.10% Indels: 0
DB: 4 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-443-184-50 (1-653)
QY 1 TGTGACCTCTGTGATCGATCATCTTTGGGATGCGCAATGGCGACGACATATAATCC 60
||||| ||| :||| ||||| ||||| |||
Db 541 CysTrpMetCysGlyLysAsnLysAsnSerGluLysGlnTrpGlnGlnHisLysSer 560
||| ||| ||| ||| ||| ||| ||| |||
QY 61 AATCCAC 69
||| |||
Db 561 GluLysHis 563

RESULT 8
US-09-288-143-190
; Sequence 190, Application US/09288143
; Patent No. 6433139
; GENERAL INFORMATION:
; APPLICANT: Brewer et al.
; TITLE OF INVENTION: 53 Human Secreted Proteins
; FILE REFERENCE: P2018P1
; CURRENT APPLICATION NUMBER: US/09/288,143

```

```

; SEQ ID NO 23
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(226)
; OTHER INFORMATION: X Is selenocysteine.
US-08-679-493A-23

Alignment Scores:
Pred. No.:      12.7          Length:      226
Score:           49.00        Matches:       8
Percent Similarity: 50.00%    Conservative: 2
Best Local Similarity: 40.00% Mismatches:     10
Query Match:      30.25%     Indels:         0
DB:                4          Gaps:           0

US-09-513-151-3_COPY_1121_1210 (1-90) x US-08-679-493A-23 (1-226)
QY      86 TTTCTTCAGTGGTGTCAGATGCGGAATTGTGATTTATGTCGCCTGCCCATTCGCGATCCCA 27
Db      87 TrpPhcCystTrpPheProrfpGlutRProSerThCysAlaIaIaglyGIgylGIuPro 106

RESULT 10
US-09-228-986-114
; Sequence 114, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; FILE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; CURRENT APPLICATION NUMBER: US/09/228, 986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 114
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-114

Alignment Scores:
Pred. No.:      19.5          Length:      762
Score:           48.50        Matches:     10
Percent Similarity: 63.16%    Conservative: 2
Best Local Similarity: 52.63% Mismatches:     6
Query Match:      29.57%     Indels:         1
DB:                4          Gaps:           1

US-09-513-151-3_COPY_1121_1210 (1-90) x US-09-228-986-114 (1-762)
QY      15 TCGATCATCATGTCGGGATCGCAATGGCGACGCGACATAAAATCCAATGCCACTT 71
Db     261 SerAsnHishStrp---SerLeuAlaGlyProValAlaIaIaIeArgValProIeu 278

RESULT 11
US-08-021-601-2
; Sequence 2, Application US/08021601
; Patent No. 5591631
; GENERAL INFORMATION:
; APPLICANT: Leppia, Stephen H.
; APPLICANT: Klumpel, Kurt R.
; APPLICANT: Nichols, Peter J.
; APPLICANT: Atora, Naveen
; TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
; FILE OF INVENTION: RELATED METHODS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Needle & Rosenberg, P.C.
```

```

STREET: 133 Carnegie Way, Suite 400
City: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/021,601
FILING DATE: 19930212
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Spratt, Gwendolyn D.,
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 1414.057
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-021-601-2

Alignment Scores:
Pred. No.:          Length:      776
Score:              Matches:     9
Percent Similarity: 48.50%
                    Conservative: 5
Best Local Similarity: 53.85%
                    Mismatches:   7
Query Match:        Indels:       5
                       Caps:       1

US-09-513-151-3-COPY_1121_1210 (1-90) x US-08-021-601-2 (1-776)
QY      24 CATTGGGAGTGCAGATG-----GGCAGCGCACATAAATCAATCCCA    68
         |||||               ::|
Db      280 HistrpsrarpserleuserclugluclyargclyleuleuylslsleughnlllePro 299
         :|||||             :::::
QY      69 CTGAACCACTGAAGA A 86
         :|||||             :::::
Db      300 IlegluProlyslYasp 305

RESULT 12
US-08-082-849B-2
; Sequence 2, Application US/08082849B
; Patent No. 5677274
; GENERAL INFORMATION:
; APPLICANT: Leppla, Stephen H.
; APPLICANT: Kimpel, Kurt R.
; APPLICANT: Alora, Naveen
; APPLICANT: Singh, Yogendra
; APPLICANT: Nichols, Peter J.
; TITLE OF INVENTION: Anthrax Toxin Fusion Proteins and
; TITLE OF INVENTION: Related Methods
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

```

      APPLICATION NUMBER: US/08/082,849B
      FILING DATE: 25-JUN-1993
      CLASSIFICATION: 514
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/021,601
      FILING DATE: 12-FEB-1993
      ATTORNEY/AGENT INFORMATION:
      NAME: Weber, Kenneth A.
      REGISTRATION NUMBER: 31,677
      REFERENCE/DOCKET NUMBER: 15280-161-1
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (415) 576-0300
      TELEFAX: (415) 576-0300
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 776 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-082-849B-2

Alignment Scores:
Pred. NO.:          19.5          Length:          776
Score:              48.50         Matches:           9
Percent Similarity: 53.85%        Conservative:    5
Best Local Similarity: 34.62%      Mismatches:     7
Query Match:         29.57%        Indels:         5
DB:                  1            Gaps:             1

US-09-513-151-3_COPY_112L_1210 (1-90) x US-08-082-849B-2 (1-776)

OY      24  CATTGGGATGCGCGANG-----GGCAGCGCACATAAATCAATCCCA 68
          |||||          ::          |||
Db      280  HSTRSPERASPSERLSEKCIUGLUGLYARGGLYLEULYSLEUGLHLEPRO 299
          :::|||||          ::|||

OY      69  CTTGAACCACTGAGAA 86
          :::|||||          ::|||
Db      300  ILEGUPTOLYSLYSASP 305

RESULT 13
PCT-US94-01624-2
Sequence 2, Application PC/TUS9401624
GENERAL INFORMATION:
APPLICANT: Leppla, Stephen H.
APPLICANT: Kimpel, Kurt R.
APPLICANT: Aroza, Naveen
APPLICANT: Singh, Yogendra
APPLICANT: Nichols, Peter J.
TITLE OF INVENTION: ANTHRAX TOXIN FUSION PROTEINS AND
TITLE OF INVENTION: RELATED METHODS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESS: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Steuart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01624
FILING DATE: June 25, 1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 15280-115
TELECOMMUNICATION INFORMATION:

```

TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 776 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-01624-2

## Alignment Scores:

Pred. No.:	Length:
Score: 19.5	776
Percent Similarity: 48.50	Matches: 9
Best Local Similarity: 53.85%	Conservative: 5
Query Match: 34.62%	Mismatches: 7
	Indels: 5
DB: 5	Gaps: 1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x PCT-US94-01624-2 (1-776)

QY 24 CATTGGGATCGCGAATG-----GGCAGCGCACATTAATCCAAATCCCA 68

DB 280 HSTRPserAspserLeuSercluglYargglYleuLeuLysLysLennllePro 299

QY 69 CTGAACCACTGAAGAA 86

DB 300 lIeGluprPolYlYsasp 305

## RESULT 14

US-09-735-934A-2

Sequence 2, Application US/09735934A

Patent No. 6372468

GENERAL INFORMATION:

APPLICANT: LI, JiaYin et al

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

FILE REFERENCE: CLO00851

CURRENT APPLICATION NUMBER: US/09/735,934A

CURRENT FILING DATE: 2000-12-14

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 878

TYPE: PRT

ORGANISM: Homo sapiens

US-09-735-934A-2

## Alignment Scores:

Pred. No.:	Length:
Score: 24	878
Percent Similarity: 48.00	Matches: 8
Best Local Similarity: 70.59%	Conservative: 4
Query Match: 47.06%	Mismatches: 5
	Indels: 0
DB: 4	Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-09-735-934A-2 (1-878)

QY 66 GGATTGGATTATGTCGCTGCCCATTCGCATCCCAATGATGATCG 16

DB 173 GlyLeuAsnTyrHISLysArgCysAlaPheSerIleProAsnAsnCysSer 189

## RESULT 15

US-09-135-782-4

Sequence 4, Application US/09135782

Patent No. 6027929

GENERAL INFORMATION:

APPLICANT: Xu, Shuang-Yong

TITLE OF INVENTION: Method For Cloning And Producing The NspI Restriction

TITLE OF INVENTION: Endonuclease In E. coli And Purification Of The

TITLE OF INVENTION: Recombinant NspI Restriction Endonuclease

FILE REFERENCE: NEB-143

CURRENT APPLICATION NUMBER: US/09/135,782

CURRENT FILING DATE: 1998-08-18  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 4  
LENGTH: 244  
TYPE: PRT  
ORGANISM: No. 6027929loc sp.  
US-09-135-782-4

## Alignment Scores:

Pred. No.:	Length:
Score: 22.4	244
Percent Similarity: 47.50	Matches: 9
Best Local Similarity: 51.85%	Conservative: 5
Query Match: 33.33%	Mismatches: 12
	Indels: 1
DB: 3	Gaps: 1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x US-09-135-782-4 (1-244)

QY 1 TGTGACCTCTGTGATGCATCATTCGCGATGGCGACGACATAAATCC 60

DB 76 CysGlnThrCysGlnGlnThrArgLeuGlyAspArg--TrpThrclYcIuIleLysSer 94

QY 61 AATCCCACTTGACCAACTG 81

DB 95 AlaAsnSerAlaSerAsnIle 101

Search completed: April 21, 2003, 18:52:19  
Job time: 2.97137 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:47:50 ; Search time 4.20225 Seconds  
(without alignments)  
8825.857 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210  
Perfect score: 164  
Sequence: 1 TGTGACCTCTGTGATGCAAT.....TGAACCACTGAGAAAAAGA 90

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xih  
-O=/cgn2.1/uspro.spool/US09513151/runat\_15042003\_141144\_26380/app.query.fasta.1.2446  
-DB=SPREMBL\_21 -OPMT=fasta -SUFFIX=n2p.rspt -MINMATCH=0.1 -IOBCL=0  
-LOOPT=0 -UNITS=dits -START=1 -END=1 -MATRIX=dlosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151 -ECGN.1.1.125 -enact.15042003\_141144\_26380 -ICPU=3  
-NO.XLPHY -NO.MMAP -LARGEOVERLY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WAIN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL\_21:.\*  
1: sp\_archaea:.\*  
2: sp\_bacteria:.\*  
3: sp\_fungi:.\*  
4: sp\_human:.\*  
5: sp\_invertebrate:.\*  
6: sp\_mammal:.\*  
7: sp\_mhc:.\*  
8: sp\_organelle:.\*  
9: sp\_phage:.\*  
10: sp\_plant:.\*  
11: sp\_rodent:.\*  
12: sp\_virus:.\*  
13: sp\_vertebrate:.\*  
14: sp\_unclassified:.\*  
15: sp\_rvirus:.\*  
16: sp\_bacteriap:.\*  
17: sp\_archaeap:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match Length	ID	Description
1	164	100.0	324	4 Q96FJ3	Q96FJ3 homo sapien

2	164	100.0	326	4	Q9NXT7	Q9NXT7 homo sapien
3	164	100.0	435	4	Q96LA5	Q96LA5 homo sapien
4	164	100.0	467	4	Q9H3H1	Q9H3H1 homo sapien
5	157	95.7	326	11	Q9DIH5	Q9DIH5 mus musculus
6	64	39.0	294	4	Q9UL40	Q9UL40 homo sapien
7	63	38.4	294	11	Q9ROB7	Q9ROB7 mus musculus
8	62	37.8	524	13	Q42147	Q42147 xenopus lae
9	61	37.2	314	5	Q9VHM6	Q9VHM6 drosophila
10	60	36.6	430	5	Q9GYG3	Q9GYG3 caenorhabdi
11	59	36.0	278	10	Q8W374	Q8W374 oryza sativ
12	58.5	35.7	439	10	Q8S9Z6	Q8S9Z6 oryza sativ
13	58	35.4	478	5	Q94540	Q94540 drosophila
14	58	35.4	578	5	Q9VA29	Q9VA29 drosophila
15	58	35.4	635	5	Q95026	Q95026 drosophila
16	58	35.4	935	5	Q9V0V6	Q9V0V6 drosophila
17	57.5	35.1	935	5	Q9V207	Q9V207 drosophila
18	57	35.2	128	12	Q8A660	Q8A660 simlan viru
19	57	34.8	818	4	Q8WU72	Q8WU72 homo sapien
20	57	34.8	837	4	Q8RTG3	Q8RTG3 homo sapien
21	57	34.8	842	4	Q9H865	Q9H865 homo sapien
22	57	34.8	902	13	Q8UWC5	Q8UWC5 gallus gall
23	57	34.8	1052	11	Q88532	Q88532 mus musculi
24	57	34.8	1057	4	Q96KX1	Q96KX1 homo sapien
25	56	34.1	346	3	Q9VHM4	Q9VHM4 drosophila
26	56	34.1	463	3	Q9S811	Q9S811 schizosacch
27	56	34.6	481	16	P72956	P72956 synechocyst
28	56	34.1	1053	4	Q95625	Q95625 homo sapien
29	56	34.1	1623	5	Q45019	Q45019 caenorhabdi
30	55.5	33.8	251	10	Q9AWK1	Q9AWK1 oryza sativ
31	55.5	33.8	336	5	Q8TOY0	Q8TOY0 apis cerana
32	55.5	33.8	539	5	Q9V4M9	Q9V4M9 drosophila
33	55.5	33.8	600	4	Q96MN7	Q96MN7 homo sapien
34	55	33.5	379	2	Q48738	Q48738 lactococcus
35	55	33.5	670	2	Q9VKS7	Q9VKS7 drosophila
36	55	34.0	670	2	Q9AEF6	Q9AEF6 leptospira
37	55	34.0	670	2	Q8VTY1	Q8VTY1 leptospira
38	55	34.0	670	2	Q8VTU9	Q8VTU9 leptospira
39	55	33.5	789	5	Q9VBX4	Q9VBX4 drosophila
40	55	33.5	3140	12	Q84925	Q84925 plum pox vi
41	54	32.9	199	4	Q96NCO	Q96NCO homo sapien
42	54	32.9	199	11	Q9CEP7	Q9CEP7 mus musculi
43	54	32.9	845	11	Q8VEH2	Q8VEH2 mus musculi
44	53	32.3	225	16	Q8ZCD7	Q8ZCD7 yersinia pe
45	53	32.7	298	17	Q9H1L4	Q9H1L4 thermoplasma

## ALIGNMENTS

RESULT 1

ID	Q96FJ3	PRELIMINARY:	PRT:	324 AA.
AC	Q96FJ3			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	Similar to crna isopenstenylyptrophosphate transferase.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=UTERUS;			
RA	Strausberg R.;			
RT	Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: BC010741; AAH10741.1;			
DR	InterPro: IPR002627; IPTT			
DR	InterPro: IPR000822; znf_C2H2.			
DR	Pfam: PF01715; IPTT: 1.			
DR	ProDom: PD004674; IPTT: 1.			
DR	SMART: SM00355; znf_C2H2; 1.			
DR	PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.			
KM	Transferase.			

SO SEQUENCE 324 AA: 37223 MW: 1E6835D7C09126A9 CRC64;

Alignment Scores:

Pred. No.: 7,31e-17 Length: 324  
Score: 164.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q96FJ3 (1-324)

QY 1 TGTGACCTCTGTGATGCAATCATCTGGGATCGGCAATGGCAGCGCATATAATCC 60

DB 254 CysAspLeucCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 273

QY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90

DB 274 LysSerHisLeuAsnGlnLeuLysArg 283

RESULT 2

Q9NXT7 PRELIMINARY: PRT; 326 AA.

AC Q9NXT7;  
DT 01-OCT-2000 (TRENBLREL. 15, Created)  
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)  
DE 01-MAR-2002 (TRENBLREL. 20, Last annotation update)  
DE CDNA FLJ20061 fis, clone COL01383.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=COLON;  
RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
RA Okitani R., Ota T., Suzuki Y., Ohashi M., Nishi T., Shidohara T.,  
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;  
RT NEDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF000068; BAA90923.1;  
DR InterPro: IPR002627; IPTT.  
DR InterPro: IPR000832; Znf\_C2H2.  
DR Pfam: PF01715; IPTT; 1.  
DR ProDom: PD004674; IPTT; 1.  
DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
SQ SEQUENCE 326 AA: 37435 MW: EA83F0F664B7ACE CRC64;

Alignment Scores:

Pred. No.: 7,31e-17 Length: 326  
Score: 164.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q9NXT7 (1-326)

QY 1 TGTGACCTCTGTGATGCAATCATCTGGGATCGGCAATGGCAGCGCATATAATCC 60

DB 256 CysAspLeucCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 275

QY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90

DB 276 LysSerHisLeuAsnGlnLeuLysArg 285

RESULT 3

Q96L45 PRELIMINARY: PRT; 435 AA.

AC Q96L45;  
DT 01-DEC-2001 (TRENBLREL. 19, Created)  
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)  
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)

DE tRNA isopentenyl transferase (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-21444833; PubMed-11560893;

RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,

RA Barnes T., Hekimi S.;

RT "Regulation of Physiological Rates in Caenorhabditis elegans by a

RT tRNA-Modifying Enzyme in the Mitochondria.";

RL Genetics 159:147-157(2001).

DR EMBL; AF052768; AAL14107.1; -

DR InterPro: IPR002627; IPTT.

DR Pfam: PF01715; IPTT; 1.

DR ProDom: PD004674; IPTT; 1.

DR SMART: SM00355; Znf\_C2H2\_1.

DR TIGRFAMs: TIGR00174; mlaa; 1.

DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.

KW transferase.

FT NON\_TER

SQ SEQUENCE 435 AA: 48948 MW: 2279AE7C2D999FF1 CRC64;

Alignment Scores:

Pred. No.: 7,4e-17 Length: 435  
Score: 164.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q96L45 (1-435)

QY 1 TGTGACCTCTGTGATGCAATCATCTGGGATCGGCAATGGCAGCGCATATAATCC 60

DB 365 CysAspLeucCysAspArgIleIleIleGlyAspArgGluTrpAlaIleHisIleLysSer 384

QY 61 AAATCCCACTTGAAACCACTGAAGAAAGA 90

DB 385 LysSerHisLeuAsnGlnLeuLysArg 394

RESULT 4

Q9H3H1 PRELIMINARY: PRT; 467 AA.

AC Q9H3H1;  
DT 01-MAR-2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)  
DE tRNA isopentenylpyrophosphate transferase.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RA MEDLINE-20564178; PubMed-11111046;

RA Golovko A., Hjalms G., Sitbon F., Nicander B.;

RT "Cloning of a human tRNA isopentenyl transferase.";

RL Gene 238:85-93(2000).

DR EMBL; AF074918; AAG31324.1; -

DR InterPro: IPR002627; IPTT.

DR InterPro: IPR000832; Znf\_C2H2.

DR Pfam: PF01715; IPTT; 1.

DR ProDom: PD004674; IPTT; 1.

DR TIGRFAMs: TIGR00174; mlaa; 1.

DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.

KW transferase.

SQ SEQUENCE 467 AA: 52725 MW: 634469919D7F56A5 CRC64;

Alignment Scores:

Pred. No.: 7,42e-17 Length: 467





```

DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE Double-stranded RNA-binding zinc finger protein JAZ.
GN ZFP346.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCB1_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99419013; PubMed=10488071.
RA Yang M., May W.S., Ito T.;
RT "JAZ requires the double-stranded RNA-binding zinc finger motifs for
RT nuclear localization."
RL J. Biol. Chem. 274:27399-27406(1999).
DR EMBL: AF083339; AAD52017.1; -.
DR MGP: MGI:1349417; Zfp346.
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR003604; Znf_U1.
DR Pfam: PF00096; ZF-C2H2; 3.
DR SMART: SM00355; ZNF_C2H2; 4.
DR SMART: SM00451; ZNF_U1; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; UNKNOWN.4.
KW DNA-binding; Zinc-finger.
SO SEQUENCE 294 AA; 32698 MW; 6BD916262EDBA71E CRC64;

Alignment Scores:
Pred. No.: 0.57 Length: 294
Score: 63.00 Matches: 10
Percent Similarity: 65.52% Conservative: 9
Best Local Similarity: 34.48% Mismatches: 10
Query Match: 38.41% Indels: 0
DB: 11 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q9R0B7 (1-294)
QY 1 TGTGACCTGTGATCGATCATTCATGCGGATCGGATGCGGACGCATATAATCC 60
DB 75 CysLysValCysCysAlaMetLeuIleSerGlnLysLeuAlaHisTyrGlnSer 94
QY 61 AAATCCCACTTGACCACTGAAGAAA 87
DB 95 LysLysHisAlaAsnLysValLysArg 103

RESULT 8
042147 PRELIMINARY: PRT: 524 AA.
AC 042147.
DT 01-JAN-1998 (TRENBLREL. 05, Created)
DT 01-JAN-1998 (TRENBLREL. 05, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE DSRBP-2Fa.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCB1_TaxID=8355.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97415754; PubMed=9268652;
RA Finerty P.J., Jr., Bass B.L.;
RT "A Xenopus zinc finger protein that specifically binds dsRNA and RNA-
RT DNA hybrids."
RL J. Mol. Biol. 271:195-208(1997).
DR EMBL: AF005083; AAC60260.1; -.
DR InterPro: IPR000822; Znf_C2H2.
DR InterPro: IPR000690; Znf_Matrin.
DR InterPro: IPR003604; Znf_U1.
DR Pfam: PF00096; ZF-C2H2; 5.
DR SMART: SM00355; ZNF_C2H2; 7.
DR SMART: SM00451; ZNF_U1; 7.
DR PROSITE: PS00028; ZINC_FINGER_C2H2_1; 7.
KW DNA-binding; Metal-binding; Zinc-finger.
SO SEQUENCE 524 AA; 55588 MW; 104D86038ADB8D57 CRC64;

```

```

Alignment Scores:
Pred. No.: 0.84 Length: 524
Score: 62.00 Matches: 8
Percent Similarity: 65.52% Conservative: 11
Best Local Similarity: 27.59% Mismatches: 10
Query Match: 37.80% Indels: 0
DB: 13 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x 042147 (1-524)
QY 1 TGTGACCTGTGATCGATCATTCATGCGGATCGGATGCGGACGCATATAATCC 60
DB 36 CysLysValCysSerAlaValLeuIleSerGlnLysLeuAlaHisTyrGlnSer 55
QY 61 AAATCCCACTTGACCACTGAAGAAA 87
DB 56 ArgLysHisAlaAsnLysValLysArg 64

RESULT 9
Q9VHM6 PRELIMINARY: PRT: 314 AA.
AC Q9VHM6.
ID Q9VHM6.
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAR-2002 (TRENBLREL. 20, Last annotation update)
DE CG11762 protein.
GN CG11762.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCB1_TaxID=7227.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.T., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K.A., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svyrskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

```

RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003680; AAF54277.1; -.  
 DR HSSP: P15822; 1BBO.  
 DR FlyBase: FBgn0037618; CG11762.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR InterPro: IPR000690; znf\_mattin.  
 DR Pfam: PF00096; zf-C2H2; 5.  
 DR SMART: SM00355; znf\_C2H2; 5.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 5.  
 DR DNA-binding: Metal-binding; zinc-finger.  
 SO SEQUENCE 314 AA; 36509 MW; C783E29C050F5D57 CRC64;

Alignment Scores:  
 Pred. No.: 1.18 Length: 314  
 Score: 61.00 Matches: 11  
 Percent Similarity: 50.00% Conservative: 4  
 Best Local Similarity: 36.67% Mismatches: 15  
 Query Match: 37.20% Indels: 0  
 DB: Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q9VHM6 (1-314)

QY 1 TGTGACCTGTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATATAAATCC 60  
 |||||:::||||||| ||| ::|||  
 Db 283 CysaspilicysaspargserpneGlnargylsalaHsleuValThrHisThrargSer 302

QY 61 AAATCCCACTGAACCACTGAGAGAAAGA 90  
 ||||| ::|||::|::|  
 Db 303 MetethisHsleuGlnasValylsGln 312

RESULT 10  
 Q9GYG3 PRELIMINARY; PRT; 430 AA.  
 AC Q9GYG3; 095UF6;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical 49.5 kDa protein (TRNA isopentenyl transferase).  
 GN ZC395.6 OR GRO-1.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA MEDLINE-99069613; PubMed-9851916;  
 RA None;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Connell M.;  
 RT "The sequence of *C. elegans* cosmid ZC395.";  
 RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA MEDLINE-2144833; PubMed-11560893;  
 RA Lemieux J., Lakowski B., Webb A., Meng Y., Ubach A., Bussiere F.,  
 RA Barnes T., Hekimi S.;  
 RT "Regulation of Physiological Rates in *Caenorhabditis elegans* by a  
 RT tRNA-Modifying Enzyme in the Mitochondria.";

RL Genetics 159:147-157(2001).  
 DR EMBL: U13642; AAG00042.2; -.  
 DR DR AIO52773; AAL14112.1; -.  
 DR InterPro: IPR002627; IPPT.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR Pfam: PF01715; IPPT; 1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
 DR Hypothetical protein; DNA-binding; Transferase; Zinc-finger.  
 SO SEQUENCE 430 AA; 49548 MW; 107E95095C81A2B9 CRC64;

Alignment Scores:  
 Pred. No.: 1.72 Length: 430  
 Score: 60.00 Matches: 9  
 Percent Similarity: 53.33% Conservative: 7  
 Best Local Similarity: 30.00% Mismatches: 14  
 Query Match: 36.59% Indels: 0  
 DB: Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q9GYG3 (1-430)

QY 1 TGTGACCTGTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATATAAATCC 60  
 |||||:::||||||| ||| |||||  
 Db 394 CysGlnilicysasnilleSerMetThrGlnylsAspaSerTPGlnLysHisileSpcl 413

QY 61 AAATCCCACTGAACCACTGAGAGAAAGA 90  
 ||||| ::|||::|::|  
 Db 414 LysLysHisLysHisHisLysLysLys 423

RESULT 11  
 Q8W374 PRELIMINARY; PRT; 278 AA.  
 AC Q8W374;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Hypothetical 29.2 kDa protein.  
 GN OSJNBAA0029C15.17.  
 OS *Oryza sativa* (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 OC NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. NIPPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,  
 RA Gansberger K., Brenner M., Burgess S., Hance M., Shwartsbeyn M.,  
 RA Tsirlin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pal G.,  
 RA Vanden S.E., Utterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.;  
 RT "Oryza sativa chromosome 10 BAC OSJNBAA0029C15 genomic sequence.";  
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC087182; AAL59020.1; -.  
 DR InterPro: IPR000822; znf\_C2H2.  
 DR InterPro: IPR003604; znf\_U1.  
 DR SMART: SM00451; znf\_U1; 2.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_2.  
 KW Hypothetical protein.  
 SO SEQUENCE 278 AA; 29241 MW; 7674AC0C23341DAC CRC64;

Alignment Scores:  
 Pred. No.: 2.42 Length: 278  
 Score: 59.00 Matches: 10  
 Percent Similarity: 57.14% Conservative: 6  
 Best Local Similarity: 35.71% Mismatches: 12  
 Query Match: 35.98% Indels: 0  
 DB: Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q8W374 (1-278)

QY 1 TGTGACCTGTGTGATCGAATCATCATTTGGGATCGCGAATGGCAGCGACATATAAATCC 60  
 ||||| ::|||::|::|  
 Db 230 CysGlnleucCysaspValleuAlaSerGlnleuAsnValAlaIleHisThrArg 249

```

QY 61 AATCCACTTGACCACTGAG 84
Db 250 LysGlnHisLeuHisArgValArg 257

RESULT 12
Q8S926 PRELIMINARY: PRT: 439 AA.
AC 08S926;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN Similar to tRNA Isopenentenyltransferase.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OJ1656.A11."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AP003448; BAB5325.1; -.
KW Transferase.
SO SEQUENCE 439 AA; 49646 MW; 74BE689EAA2EEF95 CRC64;

Alignment Scores:
Pred. No.: 2.96 Length: 439
Score: 58.50 Matches: 9
Percent Similarity: 60.00% Conservative: 9
Best Local Similarity: 30.00% Mismatches: 11
Query Match: 35.67% Indels: 1
DB: 10 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x Q8S926 (1-439)
QY 1 TGTGACCTCTGTGAT---CGAATCATCTGGGATCGCGAATGGCGACGACATAA 57
Db 399 CysGlnHisArgValArgValGlnArgGlyThrHisGluTrpGlnHisLysGln 418
QY 58 TCCAATCCCTTGACCACTGAG 87
Db 419 GlyArgCysHisArgValArgValGlnArg 428

RESULT 13
Q94540 PRELIMINARY: PRT: 558 AA.
AC 094540;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
GN Zinc-finger protein ZNF2D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97134673; PubMed=8980233;
RA Sauer F., Massaman D.A., Rubin G.M., Tjian R.;
RT "AFR11s mediate activation of transcription in the Drosophila
embryo."
RL Cell 87:1271-1284(1996);
DR EMBL, U73125; AAB40721.1; -.
DR FlyBase; FBgn0017453; ZNF2D.
DR InterPro; IPR000823; Znf_C2H2.
DR InterPro; IPR000690; Znf_matin.

```

```

DR InterPro; IPR003604; Znf_U1.
DR Pfam; PF00096; zf-C2H2; 3.
DR SMART; SM00355; ZNF_C2H2; 3.
DR SMART; SM00451; ZNF_U1; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_3.
KW DNA-binding; Zinc-finger.
SO SEQUENCE 558 AA; 61157 MW; 5DC76C1BEF2DC784 CRC64;

Alignment Scores:
Pred. No.: 3.59 Length: 558
Score: 58.00 Matches: 10
Percent Similarity: 53.57% Conservative: 5
Best Local Similarity: 35.71% Mismatches: 13
Query Match: 35.37% Indels: 0
DB: 5 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x Q94540 (1-558)
QY 1 TGTGACCTCTGTGATCGAATCATCTGGGATCGCGAATGGCGACGACATAAATCC 60
Db 178 CysGlnLeuCysAspValThrCysThrGlyThrAspAlaTyrAlaAlaHisValArgGly 197
QY 61 AATCCACTTGACCACTGAG 84
Db 198 AlaLysHisGlnLysValValLys 205

RESULT 14
Q9VA29 PRELIMINARY: PRT: 578 AA.
AC 09VA29;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN CG12071 protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis J.K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glöcker A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paigle J.M.,
RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

```

RA Spier E, Spreading A.C., Stepien M., Strong R., Sun E.  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Massarman D.A., Melnick G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveloff J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zheng F.N., Zhong W., Zhou X., Zhu S., Zhu X.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
Rt "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003773; AAF57095.1; -.  
DR HSSP; P07248; 2ADR.  
DR FlyBase; FBgn0039808; CG12071.  
DR InterPro; IPR000822; Znf.C2H2.  
DR Pfam; PF00096; zf-C2H2; 3  
DR ProDom; PD000003; Znf.C2H2; 1.  
DR SMART; SM00355; Znf\_C2H2; 3.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 3.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 3.  
KW DNA-binding; Metal-binding; zinc-finger.  
SQ SEQUENCE 578 AA; 64636 MW; 19966AD07BFA43FE CRC64;

Alignment Scores:

Pred. No.:	3.59	Length:	57.7
Score:	58.00	Matches:	11
Percent Similarity:	62.50%	Conservative:	4
Best Local Similarity:	45.83%	Mismatches:	7
Query Match:	35.37%	Indels:	2
DB:	5	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x Q9VA29 (1-578)

```

QY      1  TGTACCTCTGTGATCGAATCATCATTTGGGGATCGCGAATGGGCACGCGACATAAATCC 60
      |||:::|||||  ||:::  |||:::  ||
DB    315  CysGLuLeuCysGLyArgMetPheSerSerArgAspGluTrpSerLeuHis-----Ala 332

```

```

QY      61 AAATCCCACTTG 72
          |||||
Db      333 LysSerHisLeu 336

```

## RESULT 15

ID	Q95026	PRELIMINARY; PRT; 635 AA.
AC	Q95026;	
DT	01-DEC-2001	(TREMBLrel. 19, Created)
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)

05 *Drosophila melanogaster* (Fruit fly).  
0C Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
0C Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
0C Ephydroidea; Drosophilidae; Drosophila.  
0X NCBI\_taxid=7227;

RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY:  
RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin R., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nuncio J., Paolel J., Pargans V., Park S., Phouanevong S., Wan K.,  
RA Yu C., Lewis S.E., Rulins G.M., Celiker S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY058357; AAIJ3586.1; -;  
DR FlyBase; FBgn0017453; Zn72D.  
DR InterPro; IPR000822; znf\_C2H2.  
DR Pfam; PF00096; zf\_C2H2\_3.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWNW\_3.  
KW DNA-binding; zinc-finger.  
SQ SEQUENCE 635 AA; 66457 MW; 097515874204ECDE CRC64;

### Alignment Scores:

Pred. No.:	3.61	Length:	635
Score:	58.00	Matches:	10

Percent Similarity:	53.57%	Conservative:	5
Best Local Similarity:	35.71%	Mismatches:	13
Query Match:	35.37%	Indels:	0
DB:	5	Gaps:	0

US-09-513-151-3 COPY\_1121\_1210 (1-90) x Q95U26 (1-635)

**QY**

1 TGTGACCTCTGTGATCGAATCATCATTGGGGATCCGCATGGGCAGCGCACATAAAATCC	60
...	...     .....:

Db 255 CysGluLeuCysAspValThrCysThrGlyThrAspAlaTyrAlaAlaHisValArgGly 274

61 AAATCCCACTTGAACCAACTGAAG 84

Db 275 AlalYSHiGlnLysValLys 282

Search completed: April 21, 2003, 18:55:43  
Job time : 7.20225 secs

Job time : 7.20225 secs

Job time : 7.20225 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 18:21:20 ; Search time 2.11168 Seconds

(without alignments)  
8194.498 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Perfect score: 164

Sequence: 1 TGTGACCTCTGTGATCGAAT.....TGAACCACTGAGAAAGA 90

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ n2p.model -DEV-xlh  
-O-/cg22.1/USPTO.spool/US09513151/unal\_15042003.141144.26390/app.query.fasta.1.2446  
-DB-PIR\_73 -QFMT-fastan -SUFFIX-n2p.rpr -MINMATCH=0.1 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.col -LIST=45  
-DOCALLIGN=200 -THR.SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE-LOCAL  
-OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151.ecgnat.1.1.58.gunat.15042003.141144.26390 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGESQUEW -NEG\_SCORES=0 -MATT -LONGLOG -DEV\_TIMCOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	36.6	433	2	T27538
2	59	36.0	428	2	S67176
3	58.5	35.7	453	2	B69504
4	58	35.4	301	2	S56211
5	57	35.2	180	2	I48129
6	57	35.2	180	2	I84689
7	57	34.8	1052	2	T14343
8	57	35.2	1560	2	T54361
9	56	34.1	263	2	G87721
10	56	34.1	463	2	T41390
11	56	34.6	481	2	S74934
12	55	33.5	379	2	C55205
13	55	33.5	3140	2	S47508
14	54	32.9	222	2	H64422

15	53.5	32.6	232	2	H84028
16	53	32.3	225	2	AE0371
17	53	32.3	433	2	B29345
18	52	31.7	3140	1	GNV5RA
19	51.5	31.8	63	2	E97052
20	51.5	31.4	1044	1	DVBYE3
21	51	31.5	325	2	C90282
22	51	31.1	407	2	S40990
23	51	31.1	489	2	S44909
24	51	31.1	588	2	JC7206
25	50	30.5	104	2	S74901
26	50	30.5	208	2	D96543
27	50	30.5	257	2	B45061
28	50	30.5	260	2	A45061
29	50	30.5	367	2	AE3582
30	50	30.5	386	2	F83994
31	50	30.5	514	2	C64154
32	50	30.9	548	2	C64438
33	50	30.5	574	2	T00245
34	50	30.5	3624	2	AD0835
35	49.5	30.2	113	2	H84146
36	49.5	30.6	172	2	T44356
37	49.5	30.2	214	2	B72377
38	49.5	30.2	340	2	E95964
39	49.5	30.6	479	2	T05588
40	49.5	30.2	682	2	JC7670
41	49.5	30.2	1817	2	H71611
42	49	29.9	61	2	T41859
43	49	29.9	157	2	E96832
44	49	29.9	313	2	F64433
45	49	29.9	395	2	T52423

#### ALIGNMENTS

##### RESULT 1

T27538  
hypothetical protein ZC395.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27538

R:Connell, M. the EMBL Data Library, August 1994

A:Description: The sequence of C. elegans cosmid ZC395.

A:Reference number: Z20385

A:Accession: T27538

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-433 <CON>

A:Cross-references: EMBL:U13642; PIDN:AA53886.1; GSPDB:GN00021; CESP:ZC395.6

A:Experimental source: strain Bristol N2; clone ZC395

C:Genetics:

A:Gene: CESP:ZC395.6

A:Map position: 3

A:Introns: 43/1; 201/1; 266/3; 310/2; 341/3; 377/3; 410/2

##### Alignment Scores:

Pred. No.: 1.21  
Score: 60.00  
Percent Similarity: 53.38  
Best Local Similarity: 30.008  
Query Match: 36.598

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x T27538 (1-433)

QY	1	TGTGACCTCTGTGATCGAATCATTCGGGATCGCGAATGGCCACCATTAATCC	60
DB	397	CysGluIleCysAsnIleSerMetTrgGlyAspAsnTrpIleGlnLysAspGly	416
QY	61	AAATCCCACTTGACCAACGACGAGAAAGA	90
DB	417	LysLysHisLysHisHisLysHisLysGlnLys	426

DNA repair protein  
probable carboxype  
steroid hormone re  
genome polypeptide  
hypothetical prote  
translational elonga  
hypothetical prote  
hypothetical prote  
ZK686.4 protein -  
phosphoprotein pho  
arsenical resistanc  
unknown protein [1  
granzyme A (EC 3.4  
granzyme A (EC 3.4  
sarcosine oxidase  
xylose operon tran  
hypothetical prote  
phenylalanine-tRNA  
reverse transcript  
large repetitive P  
hypothetical prote  
probable dCTP deam  
transcription regu  
probable iron upta  
cellulase (EC 3.2.  
cathepsin B mRNA 3  
probable secreted  
ACKNPV orf122 - Bo  
hypothetical prote  
hypothetical prote  
CH2 zinc finger P

## RESULT 2

S67176 tRNA isopentenyltransferase (EC 2.5.1.8) - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein O547w; protein YOR274w  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
C:Accession: S67176; A26717; S72045  
R:Cheret, G.; Sor, F.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67169  
A:Accession: S67176  
A:Molecule type: DNA  
A:Residues: 1-428 <CH>  
A:Cross-references: EMBL:275182; NID:g1420613; PID:e252418; PID:g1420614; MIPS:YOR274w  
A:Experimental source: strain S288C  
R:Najarian, D.; Dhanich, M.E.; Martin, N.C.; Hopper, A.K.  
Mol. Cell. Biol. 7, 185-191, 1987  
A:Title: DNA sequence and transcript mapping of MOD5, features of the 5' region which su  
A:Reference number: A26717; MUID:8712703; PMID:3031457  
A:Accession: S72045  
A:Molecule type: DNA  
A:Residues: 1-374, R', 376-428 <NA>  
A:Cross-references: EMBL:M15991  
R:Cheret, G.; Bernardi, A.; Sor, F.  
Yeast 12, 1059-1064, 1996  
A:Title: DNA sequence analysis of the VP1-SNF2 region on chromosome XV of *Saccharomyces*  
A:Reference number: S72039; MUID:97051594; PMID:8896271  
A:Accession: S72045  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-428 <CH>  
A:Cross-references: EMBL:X89633; NID:g1279694; PIDN:CAA61780.1; PID:g1419759  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
C:Genetics:  
A:Gene: SGD:MOD5  
A:Cross-references: MIPS:YOR274w; SGD:S0005800  
A:Map position: 15R  
C:Keywords: transferase

## Alignment Scores:

Pred. No.:	1 72	Length:	428
Score:	59.00	Matches:	10
Percent Similarity:	57.14%	Conservative:	10
Best Local Similarity:	28.57%	Mismatches:	9
Query Match:	35.98%	Indels:	6
DB:	2	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x S67176 (1-428)

QY 1 TGTGACCTCTGT-----GATGCAATCATTCATGGGATCGCAATGG 42  
|||||:|||||  
DB 375 CysAsnValCysArgAsnAlaAspGlyLysAsnValAlaAlaIleGlyGluYrTTP 394

QY 43 GCAGCCACATTAATCCAAATCCACTGACCAACGAAGA 87  
|||||:|||||  
DB 395 LysIleHisLeuGlySerArgHisLysSerAsnLeuLysArg 409

## RESULT 3

S69504 seryl-tRNA synthetase (serS) homolog - *Archaeoglobus fulgidus*  
C:Species: *Archaeoglobus fulgidus*  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Jul-1999  
C:Accession: B69504  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sultion, G.G.; Gill, S.; Kirkness, E.F.  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Mature 350, 364-370, 1997  
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Arliah, P.; Kaine, B.P.; Sykes, S.  
Smith, H.O.; Moose, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: B69504

A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-453 <KL>  
A:Cross-references: GB:AE000962; GB:AE000782; NID:g2689285; PIDN:AAB69219.1; PID:g26  
C:Superfamily: serine-tRNA ligase

## Alignment Scores:

Pred. No.:	2 06	Length:	453
Score:	58.50 <td>Matches:</td> <td>13</td>	Matches:	13
Percent Similarity:	62.07% <td>Conservative:</td> <td>5</td>	Conservative:	5
Best Local Similarity:	44.83% <td>Mismatches:</td> <td>8</td>	Mismatches:	8
Query Match:	35.67% <td>Indels:</td> <td>3</td>	Indels:	3
DB:	2	Gaps:	1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x B69504 (1-453)

QY 4 GACCTGTGATGCAATCATTCATGGGATCGCAATGAGCAGCACAATAATCCAA 63  
|||||:|||||  
DB 27 AspIleValAspArgAlaIleGluLeuAspArgLysTrpArgLysGluLeuLys----- 44

QY 64 TCCCACTGACCACTGAAGAAGA 90  
|||||:|||||  
DB 45 ---ArgValAsnGlnLeuArgLysArg 52

## RESULT 4

S56211 probable membrane protein YPL044c - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*  
C>Date: 02-Sep-1995 #sequence\_revision 19-Oct-1995 #text\_change 19-Apr-2002  
C:Accession: S56211  
R:Murakami, Y.; Naitou, M.; Hagihara, H.; Shibata, T.; Ozawa, M.; Sasana, S.I.; Sa  
submitted to the EMBL Data Library, May 1995

A:Description: Analysis of the nucleotide sequence of chromosome VI from *Saccharomyces*  
A:Reference number: S56186  
A:Accession: S56211  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-301 <MUR>  
A:Cross-references: EMBL:D50617; NID:g836685; PID:d1009835; PID:g836711; MIPS:YPL044

C:Genetics:  
A:Cross-references: SGD:S0001850  
A:Map position: 6L  
C:Keywords: transmembrane protein  
F:179-195/Domain: transmembrane #status predicted <TRM>

Alignment Scores:

Pred. No.:	2 42	Length:	301
Score:	58.00 <td>Matches:</td> <td>9</td>	Matches:	9
Percent Similarity:	60.00% <td>Conservative:</td> <td>6</td>	Conservative:	6
Best Local Similarity:	36.00% <td>Mismatches:</td> <td>10</td>	Mismatches:	10
Query Match:	35.37% <td>Indels:</td> <td>0</td>	Indels:	0
DB:	2	Gaps:	0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x S56211 (1-301)

QY 1 TGTGACCTGTGATGCAATCATTCATGGGATCGCAATGAGCAGCACAATAATCC 60  
|||||:|||||  
DB 272 CysAsnThrCysGlnMetThrPheValGlyGluArgValAlaAlaArgHisAlaGluSer 291

QY 61 AAATCCCACTGAAC 75  
|||||:|||||  
DB 292 ThrGlyHisValAsp 296

## RESULT 5

I48129 x6169 (escapes X-inactivation) protein - hamster (fragment)

C:Species: *Cricetinae* gen. sp. (hamster)

C>Date: 12-Aug-1996 #sequence\_revision 12-Aug-1996 #text\_change 26-May-2000  
C:Accession: I48129

R.Wu, J. Salido, E.C.; Yen, P.H.; Mohandas, T.K.; Heng, H.H.; Tsui, L.C.; Park, J.;  
Mature Genet. 7, 491-496, 1994  
A:Title: The murine x6169 gene escapes X-inactivation like its human homologue.

A:Reference number: I48129; MUID:95038834; PMID:7951318

[illegible]

Db 1234 AlATrPrpLUTrPaspRhrLysPheLeuCys-----ProLeuCys 1247

OY 20 ATTCGATCAGAGG 6

Db 1248 MetArSerArGArG 1252

# RESULT 9

protein zc123.3 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001

C/Accession: G87721

R/Anonymous: The C. elegans Sequencing Consortium.

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A/Reference number: A75000: MUID:99069613; PMID:9851916

A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/

A/Note: published errata appeared in Science 283, 35, 1999; Science 285, 2103, 1999; and

A/Accession: G87721

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-263 <STO>

A/Cross-references: GB:chr.I; PIDN:AB97603.1; PID:g2804499; GSPDB:GN00019; CESP:ZC123.3

A/Note: contains similarity to C2H2-type zinc fingers

C/Genetics:

A/Gene: ZC123.3

A/Map position: 1

Alignment Scores:

Pred. No.: 4.88 Length: 263

Score: 56.00 Matches: 9

Percent Similarity: 50.00% Conservative: 6

Best Local Similarity: 30.00% Mismatches: 15

Query Match: 34.15% Indels: 0

DB: 2 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x G87721 (1-263)

OY 1 TGTGACCTCTGTGATTCATCATTCGGGATGCGGATGGCGACATATAAATCC 60

Db 204 CysGlnuMetCysSerGluThrPheAlaThrLysGlnuAlaPheLeuSerHisLeuAsnSer 223

OY 61 AATATCCACTGTAACCACTGAGAAAGA 90

Db 224 AlaArgHisLeuGlnGlnAlaLysGln 233

# RESULT 10

T41390

zinc finger protein - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x T41390 (1-463)

OY 1 TGTGACCTCTGTGATTCATCATTCGGGATGCGGATGGCGACATATAAATCC 60

Db 72 CysGlnuValCysAsnLysPheTyrSerGluLysAlaTyrSerHisMetAlaSer 91

OY 61 AATATCCACTGTAACCACTGAGAAAGA 87

Db 92 LysLysHisArgAspAsnLeuSerLys 100

# RESULT 11

574934

pled-1 protein - Synechocystis sp. (strain PCC 6803)

N/Alternate names: protein slr0687

C/Species: Synechocystis sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C/Accession: S74934

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O.; K.; Okumura, S.; Shimp, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ye

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A/Reference number: S74934; MUID:97061201; PMID:8905231

A/Accession: S74934

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-481 <KAN>

A/Cross-references: EMBL:D90902; GB:AB01339; NID:91652027; PIDN:BA16974.1; PID:dlc

C/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Genetics:

A/Gene: pled-1

A/Start codon: GTG

C/Superfamily: response regulator homology

C/Keywords: phosphoprotein

F:10-122/Domain: response regulator homology <RRH>

F:61/Binding site: phosphate (Asp) (covalent) #status predicted

Alignment Scores:

Pred. No.: 5.01 Length: 481

Score: 56.00 Matches: 10

Percent Similarity: 61.54% Conservative: 6

Best Local Similarity: 38.46% Mismatches: 8

Query Match: 34.57% Indels: 2

DB: 2 Gaps: 1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x S74934 (1-481)

OY 80 AGTGTCAAGTGAGATTGATTTATGCGCTGCCATTCGCGATCCCAATGATG 21

Db 281 SerTrpLeuProTrpGluLysGlnPheLeuCys-----HisLeuSerSerGlnValAla 298

OY 20 ATTCGATCAGAGGTCA 3

Db 299 IleAlaIleGlnGlnSer 304

# RESULT 12

C55205

Integrase - Lactococcus lactis transposon Tn5276

C/Species: Lactococcus lactis

C/Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 15-Oct-1999

C/Accession: C55205

R/Rauch, P.J.G.; de Vos, W.M.

J. Bacteriol. 176, 2165-2171, 1994

A/Title: Identification and characterization of genes involved in excision of the Ia

A/Reference number: A55205; MUID:94209213; PMID:8157585

A/Accession: C55205

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-379 <RAU>

A/Cross-references: GB:U27649; NID:9497770; PIDN:AAA7427.1; PID:9497773

C/Genetics:





Db 208 GluLeuAspHisValIleIleGlyAspArgTrpPheValSer---LeuLysGluLys 226  
Oy 64 TCCCACTG 72  
          |||||  
Db 227 GlyHisLeu 229

Search completed: April 21, 2003, 18:51:12  
Job time : 5.11168 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: April 21, 2003, 17:14:14 ; Search time 1.07696 Seconds  
(without alignments)  
6932.240 Million cell updates/sec

Title: US-09-513-151-3\_COPY\_1121\_1210

Sequence: 1 TGTGACCTCGTGTGATCGAAT.....TGACCACTGAGAGAAAGA 90

Scoring table:

BIOSUM62	
Xgapop 10.0, Xgapext 0.5	
Xgapop 10.0, Xgapext 0.5	
Fgapop 6.0, Fgapext 7.0	
Delop 6.0, Delext 7.0	

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+np.model -DEV=xln  
-O/cgn2\_1/USPTO.spool/US09513151/unat\_15042003\_141143\_26366/app\_query.fasta.1.2446  
-DB=SwissProt.40 -QFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=200000000  
-USER=US09513151 -ECGN\_1\_1\_25=unat\_15042003\_141143\_26366 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAR -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt.40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Length	ID	Description
1	63	38.4	845	1 MAT3_RAT
2	63	38.4	847	1 MAT3_HUMAN
3	59	36.0	428	1 MOD5_YEAST
4	58.5	35.7	453	1 SYS_ARCFU
5	58	35.4	301	1 YFE4_YEAST
6	57	35.2	180	1 SMCX_YEAST
7	57	34.8	898	1 C121_HUMAN
8	57	35.2	1539	1 SMCY_HUMAN
9	57	35.2	1548	1 SMCY_MOUSE
10	57	35.2	1554	1 SMCX_MOUSE
11	57	35.2	1560	1 SMCX_HUMAN
12	54	32.9	222	1 T2M1_METJA
13	54	32.9	500	1 ERR2_HUMAN
14	53.5	32.6	232	1 ERR2_HUMAN
15	53	32.3	433	1 ERR2_RAT
16	53	32.3	458	1 ERR3_HUMAN
17	52	31.7	3140	1 POLG_PPVRA
18	52	31.7	3140	1 POLG_PPVSK

19	51.5	31.4	1043	1 EFA3_YEAST
20	51	31.1	350	1 RRB3_RHISN
21	51	31.1	390	1 YMF7_CAEBL
22	51	31.1	407	1 Y014_CAEBL
23	51	31.5	507	1 SEPT1_MOUSE
24	51	31.1	845	1 KRUI_DROME
25	50	30.5	116	1 ZTR6_HUMAN
26	50	30.5	116	1 ZTR6_MOUSE
27	50	30.5	260	1 GRA_MOUSE
28	50	30.5	514	1 YUJI_HAETN
29	50	30.5	548	1 SYEB_METJA
30	50	30.5	717	1 PRDD_HUMAN
31	50	30.5	733	1 YV34_HUMAN
32	50	30.9	993	1 ROXN_HUMAN
33	49.5	30.6	172	1 DCD_CLOTH
34	49.5	30.2	961	1 BASO_MOUSE
35	49	29.9	313	1 YV71_METJA
36	49	29.9	993	1 ROXN_HUMAN
37	48.5	29.6	809	1 LER_BACAN
38	48	29.3	244	1 LPR4_MYCTU
39	48	29.3	295	1 MIRA_CAUCR
40	48	29.3	332	1 PDXA_FUSNN
41	48	29.3	456	1 CN7A_MOUSE
42	48	29.6	457	1 IREP_MOUSE
43	48	29.6	460	1 NUD4M_SCYCA
44	48	29.3	638	1 YN32_CAEBL
45	48	29.3	646	1 N935_HUMAN

## ALIGNMENTS

RESULT 1  
MAT3\_RAT  
ID MAT3\_RAT STANDARD: PRT: 845 AA.  
AC P43244; 035833;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Matrix 3.  
GN MATR3.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=91236771; PubMed=2033075;  
RA Belgrader P.; Dev R.; Berezney R.;  
RT "Molecular cloning of matrin 3, A 125-kilodalton protein of the  
RT nuclear matrix contains an extensive acidic domain.";  
RT J. Biol. Chem. 266:9893-9899(1991).  
RN [2]  
RP REVISIONS.  
RA Berezney R.;  
RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
RT -1- FUNCTION: MAY PLAY A ROLE IN TRANSCRIPTION OR MAY INTERACT WITH  
RT OTHER NUCLEAR MATRIX PROTEINS TO FORM THE INTERNAL FIBROGRAMULAR  
RT NETWORK.  
RT -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX.  
RT -1- SIMILARITY: CONTRAINS 1 MATRIN-TYPE ZINC FINGER.  
RT -1- SIMILARITY: CONTRAINS 2 RNA RECOGNITION MOTIFS (RNM).  
RT This SWISS-PROT entry is copyright. It is produced through a collaboration  
RT between the Swiss Institute of Bioinformatics and the EMBL outstation -  
RT the European Bioinformatics Institute. There are no restrictions on its  
RT use by non-profit institutions as long as its content is in no way  
RT modified and this statement is not removed. Usage by and for commercial  
RT entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
RT or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; M63485; AAB63955.1; -;  
DR InterPro; IPR000504; RNA\_rec\_mot.

```

DR InterPro: IPRO003604; znf_U1.
DR InterPro: IPRO000690; znf_matriin.
DR Pfam: PF00076; trm_2.
DR SMART: SM00360; RRM_2.
DR SMART: SM00355; znf_C2H2_1.
DR SMART: SM00451; znf_U1_2.
DR PROSITE: PSS0102; RRM_2.
DR PROSITE: PSS0030; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PSS0171; ZF_MATIN; 1.
DR Nuclear protein; RNA-binding. Repeat; Zinc-finger.
FT DOMAIN 398 473
FT DOMAIN 496 571 RNA-BINDING (RRM) 1.
FT DOMAIN 708 716 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT ZN_FING 799 830 MATRIX-TYPE.
SQ SEQUENCE 845 AA; 94446 MW; 5EA8D05529171238 CRC64;

Alignment Scores:
Pred. No.: 0.154 Length: 845
Score: 63.00 Matches: 9
Percent Similarity: 60.87% Conservative: 5
Best local Similarity: 39.13% Mismatches: 9
Query Match: 38.41% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x MAT3_RAT (1-845)
QY 1 TGTGACCTGCTGATGATCAATCATCATTTGGGAGTCCGATGGGACACCATAAATCC 60
Db 233 CysSerIleCysAspLeuProValHisSerAsnIlyGluTrpSerIleHisIleAsnGly 312
QY 61 AAATCCAC 69
Db 313 AlaserHis 315

RESULT 2
MAT3_HUMAN
ID MAT3_HUMAN STANDARD; PRT; 847 AA.
AC P43243; Q9UQ27; Q9UHW0;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Matriin 3.
GN MATR3 OR KIAA0723.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid:9606;
XP 1
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=99087487; PubMed=9872452;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. XI.
RA The complete sequences of 100 new cDNA clones from brain which code
RA for large proteins in vitro.";
RL DNA Res. 5:277-286(1998).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Hypothalamus;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Kong R., Ye M., Zhou J.,
RA Xu S.-H., Gu Z., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RA "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RA axis and full-length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).

[3]
RP SEQUENCE OF 429-847 FROM N.A.
RX MEDLINE=9126771; PubMed=2033075;
RA Belgrader P., Dey R., Bereznay R.;

```

[illegible]

RESULT 3  
MOD5\_YEAST STANDARD: PRT: 428 AA.  
AC P07864: 012203;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE tRNA isopentenyltransferase (EC 2.5.1.8) (Isopentenyl-diphosphate: DE tRNA isopentenyltransferase) (IPTase) (IPTase) (IPTase).  
GN MOD5 OR YOR0274W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE-87172703; PubMed-3031457;  
RA Najarian D., Dhanich M.E., Martin N.C., Hopper A.K.;  
RT "DNA sequence and transcript mapping of MOD5: features of the 5' RT region which suggest two translational starts.";  
RL Mol. Cell. Biol. 7:185-191(1987).  
[2]  
RN  
RP SEQUENCE FROM N.A.  
RX STRAIN-S288C;  
RX MEDLINE-97051594; PubMed-8896271;  
RA Cheret G., Bernardi A., Sor F.J.;  
RT "DNA sequence analysis of the VP1-SNF2 region on chromosome XV of RL Saccharomyces cerevisiae.";  
RL Yeast 12:1059-1064(1996).  
[3]  
RN  
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.  
RX MEDLINE-92052176; PubMed-1946403;  
RA Slusher L.B., Gillman E.C., Martin N.C., Hopper A.K.;  
RT "mRNA leader length and initiation codon context determine RT alternative AUG selection for the yeast gene MOD5.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9789-9793(1991).  
[4]  
RN  
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.  
RX MEDLINE-91203856; PubMed-1850093;  
RA Gillman E.C., Slusher L.B., Martin N.C., Hopper A.K.;  
RT "MOD5 translation initiation sites determine N6-isopentenyladenosine RT modification of mitochondrial and cytoplasmic tRNA.";  
RL Mol. Cell. Biol. 11:2382-2390(1991).  
[5]  
RN  
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.  
RX MEDLINE-94187700; PubMed-8139535;  
RA Boguta M., Hunter L.A., Shen W.C., Gillman E.C., Martin N.C., RT Hopper A.K.;  
RT "Subcellular locations of MOD5 proteins: mapping of sequences RT sufficient for targeting to mitochondria and demonstration that RT mitochondrial and nuclear isoforms comingle in the cytosol.";  
RL Mol. Cell. Biol. 14:2298-2306(1994).  
CC  
CC -1- FUNCTION: RESPONSIBLE FOR THE MODIFICATION OF A37 TO ISOPENTENYL CC A37 OF BOTH CYTOSOLIC AND MITOCHONDRIAL TRNAS.  
CC  
CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA = diphosphate + CC tRNA containing 6-isopentenyladenosine.  
CC  
CC -1- PATHWAY: BIOSYNTHESIS OF THE MODIFIED BASE ISOPENTENYLADENOSINE CC IN TRNAS.  
CC  
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL, CYTOPLASMIC AND NUCLEAR.  
CC  
CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.  
CC  
CC -1- SIMILARITY: BELONGS TO THE IPT TRANSFERASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC between the Swiss Institute of Bioinformatics and the EMBL outstation- CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (see http://www.isb-slb.ch/announce/ CC or send an email to license@isb-slb.ch).  
CC  
CC EMBL; M15991; AAA34785.1; -

DR EMBL; X89633; CAA61780.1; -  
DR EMBL; Z75182; CAA9499.1; -  
DR PIR; A26717; A26717.  
DR SGD; S0005800; MOD5.  
DR InterPro: IPR002627; IPT.  
DR Pfam; PF01715; IPT; 1.  
DR ProDom; PD004674; IPT; 1.  
DR TIGRFAMs; TIGR00174; miaA; 1.  
KW Transferase; tRNA processing; ATP-binding; Alternative Initiation;  
KW Mitochondrion; Nuclear protein.  
FT CHAIN 1 428  
FT  
FT CHAIN 12 428  
FT  
FT INIT\_MET 12 12  
FT NP\_BIND 21 28  
FT DOMAIN 210 232  
FT  
FT CONFLICT 313 313  
FT CONFLICT 375 375  
SQ SEQUENCE 428 AA; 50236 MW; A956B17ABC05161F CRC64;  
  
Alignment Scores:  
Pred. No.: 0.616 Length: 428  
Score: 59.00 Matches: 10  
Percent Similarity: 57.14% Conservative: 10  
Best Local Similarity: 28.57% Mismatches: 9  
Query Match: 35.98% Indels: 6  
DB: 1 Gaps: 1  
  
US-09-513-151-3\_COPY\_1121\_1210 (1-90) x MOD5\_YEAST (1-428)  
QY 1 TGTCACCTCTGTG-----GATCGAATCATCATGCGGATCGCATG 42  
|||:|||||:|  
Db 375 CysAsnValCysArgAsnAlaAspCylLysAsnValAlaIleGlcLulysYtrp 394  
QY 43 GCAGGCACATTAATCCAAATCCCATGTCGACCACTGACCACTGAAGAA 87  
|||:|||||:|  
Db 395 LysIleHsLsleuGlySerArgArgHisLysSerAsnLeuLysArg 409  
  
RESULT 4  
SYS\_ARCFU STANDARD: PRT: 453 AA.  
ID SYS\_ARCFU  
AC 028244;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Seryl-tRNA synthetase (EC 6.1.1.11) (Seryl-tRNA ligase) (SERs).  
GN SER5 OR AF2035.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
OC Archaeoglobaceae; Archaeoglobus.  
OX NCBI\_TaxID=2234;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RX STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE-98049343; PubMed-9389475;  
RA Kleck H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gynn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Klotzness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,  
RA Sadov P.W., D'Andrea K.P., Bowman C., Fujil C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).  
CC  
CC -1- CATALYTIC ACTIVITY: ATP + L-serine + tRNA(Ser) = AMP + diphosphate  
CC + L-seryl-tRNA(Ser).

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000962; AAB89219.1; -
CC HSSP: P34945; 1SES.
CC TIGR: AF2035; -
CC InterPro: IPR002106; AAtRNA_ligase1.
CC InterPro: IPR002314; tRNA-synt_2b.
CC InterPro: IPR002317; tRNA-synt_ser.
CC Pfam: PF00587; tRNA-synt_2b; 1.
CC Pfam: PF02403; SerYLtRNA.N; 1.
CC PRINTS: PR00981; TRNASYNTHSER.
CC TIGRfams: TIGR00414; serS; 1.
CC DR PROSITE: PS50862; AA_tRNA_LIGASE_II; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC complete proteome.
CC SEQUENCE 453 AA; 52455 MW; 4BCE9EB4F606D90 CRC64;

Alignment Scores:
Pred. No.: 0.739 Length: 453
Score: 58.50 Matches: 13
Percent Similarity: 62.07% Conservative: 5
Best Local Similarity: 44.83% Mismatches: 8
Query Match: 35.67% Indels: 3
DB: 1 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x SYS_ARCFU (1-453)
QY 4 GACCTCTGTGATCATCATTCATTCGGATCGGATCGGACGACATTAATCAAA 63
DB 27 AapllleValAspArgAlaIleGluLeuAspArgLysTrpArgGluGluLeuLys----- 44
QY 64 TCCCACTTGACCAACGACGAGAAAGA 90
DB 45 ---ArgValAsnGlnLeuArgLysArg 52

RESULT 5
YFE4_YEAST
ID YFE4_YEAST STANDARD: PRT; 301 AA.
AC P43558;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 33.5 kDa protein in SEC53-FET5 Intergenic region.
GN YFL044C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murkaml Y., Natou M., Haglawa H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Seda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae."
RT Nat. Genet. 10:261-268(1995).
CC -1- SIMILARITY: TO S.POMBE SPAC24C9.14.
CC -1- SIMILARITY: CONTAINS 1 OTU DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D50617; BAA09197.1; -
CC SGD: S0001850; YFL044C.
CC InterPro: IPR003323; OTU.
CC Pfam: PF02338; OTU; 1.
CC SMART: SM00355; ZnF_C2H2; 1.
CC DR PROSITE: PS50802; OTU; 1.
CC KW Hypothetical protein.
CC FT DOMAIN 109 OTU.
CC SEQUENCE 301 AA; 33510 MW; 32ABA3CFEA4B5EB CRC64;

Alignment Scores:
Pred. No.: 0.86 Length: 301
Score: 58.00 Matches: 9
Percent Similarity: 60.00% Conservative: 6
Best Local Similarity: 36.00% Mismatches: 10
Query Match: 35.37% Indels: 0
DB: 1 Gaps: 0

US-09-513-151-3_COPY_1121_1210 (1-90) x YFE4_YEAST (1-301)
QY 1 TGTGACCTCTGTGATCATCATTCATTCGGATCGGACGACGACATTAATCC 60
DB 272 CysAsnThrCysGlnMetThrPheValGlyGluArgLysValAlaArgHisAlaGluSer 291
QY 61 AATCCCACTTGAC 75
DB 292 ThrGlyHisValAsp 296

RESULT 6
SMCX_CRIGR
ID SMCX_CRIGR STANDARD: PRT; 180 AA.
AC P41228;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE SMCX protein (Xe169 protein) (fragment).
GN SMCX OR Xe169.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95038834; PubMed=7951318;
RA Wu J., Salido E., Yen P., Mohandas T., Shapiro L.J.;
RT "The murine Xe169 gene escapes X-inactivation like its human
RT homologue."
RT Nat. Genet. 7:491-496(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L29564; AAA62383.1; -
CC FT NON_TER 1
CC FT NON_TER 180
CC SEQUENCE 180 AA; 19964 MW; 89D2A3C50B4A703A CRC64;

Alignment Scores:
Pred. No.: 1.19 Length: 180
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3

```







```

CC -1- SIMILARITY: CONTAINS 1 JMUN DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF127244; AAD53048.1;
DR EMBL: Z29652; CAA82760.1;
DR MGD: MGI:99780; Smcy.
DR InterPro: IPR001606; ARID.
DR InterPro: IPR003347; TE_JmJc.
DR InterPro: IPR003349; TE_JmJN.
DR InterPro: IPR004198; Znf_C5HC2.
DR InterPro: IPR001965; Znf_PHD.
DR Pfam: PF00628; PHD; 4.
DR Pfam: PF01388; ARID; 2.
DR Pfam: PF02373; JmJc; 2.
DR Pfam: PF02375; JmJN; 2.
DR Pfam: PF02928; Zf-C5HC2; 1.
DR SMART: SM00501; BRIGHT; 1.
DR SMART: SM00545; JmJN; 1.
DR SMART: SM00249; PHD; 2.
DR PROSITE: PS01359; ZF-PHD_1; 2.
DR PROSITE: PS50016; ZF-PHD_2; 1.
DR ZINC-finger; Repeat.
FT DOMAIN 13 59 JMUN.
FT ZN_FING 76 184 ARID.
FT DOMAIN 325 371 PHD-TYPE 1.
FT ZN_FING 500 616 JmJc.
FT ZN_FING 1182 1243 PHD-TYPE 2.
FT CONFLICT 160 162 ILY -> SFT (IN REF. 2).
FT CONFLICT 167 167 F -> S (IN REF. 2).
FT CONFLICT 215 215 R -> K (IN REF. 2).
FT CONFLICT 583 586 VRIN -> STIK (IN REF. 2).
SQ SEQUENCE 1548 AA; 177016 MW; 44AAB94708EAA402 CRC64;

Alignment Scores:
Pred. No.: 1.37 Length: 1548
Score: 57.00 Matches: 9
Percent Similarity: 64.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 3
Query Match: 35.19% Indels: 6
DB: 1 Gaps: 1

US-09-513-151-3_COPY_1121_1210 (1-90) x SMCY_MOUSE (1-1548)
QY 80 AGTGGTTCAGAGTGCGATTGATTTATGCGCGCTGCCCATTCGCATCCCAATGATG 21
Db 1229 Alatrptgltutrpaspfthlyshphelucys-----Proleucys 1242
QY 20 ATTGCATCACAGAG 6
Db 1243 Metatgseratgarg 1247

RESULT 10
SMCX_MOUSE STANDARD; PRT: 1554 AA.
AC PA1230; Q8CIV4; Q9DDC3; Q9QVR8; Q9R039; O54995;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SMCX protein (Xel69 protein).
GN SMCX OR XE169.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN-129/SyJ;
RA MEDLINE=99373259; PubMed=10441747;
RX Agulnik A.I., Longepied G., Ty M.T., Bishop C.E., Mitchell M.J.;
RT "Mouse H-Y encoding Smcy gene and its X chromosomal homolog smcx.";
RL Mamm. Genome 10:926-929(1999).
[12]
RP SEQUENCE OF 36-1068 FROM N.A.
RC STRAIN-BALB/c; TISSUE=Testis;
RX MEDLINE=95038739; PubMed=7951230;
RA Agulnik A.I., Mitchell M.J., Matzel M.-G., Borsani G., Avner P.A.,
RT Lerner J.L., Bishop C.E.;
RP "A novel X gene with a widely transcribed Y-linked homologue escapes
RT X-inactivation in mouse and human.";
RL Hum. Mol. Genet. 3:879-884(1994).
[13]
RP SEQUENCE OF 1206-1385 FROM N.A. (ISOFORM 1).
RX MEDLINE=95038834; PubMed=7951318;
RA Wu J., Salido E., Yen P., Mohandas T., Shapiro L.J.;
RT "The murine Xel69 gene escapes X-inactivation like its human
RT homologue.";
RL Nat. Genet. 7:491-496(1994).
[14]
RP SEQUENCE OF 1-258 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo, and Small Intestine;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai T., Shunagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov M., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirini L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Butt C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszawski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[15]
RP SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE=98389008; PubMed=9723615;
RA Jegalian K.G., Page D.C.;
RT "A proposed path by which genes common to mammalian X and Y
RT chromosomes evolve to become X inactivated.";
RL Nature 394:776-780(1998).
[16]
RP FUNCTION.
RC STRAIN=C3H;
RX MEDLINE=95379952; PubMed=7544442;
RA Scott D.M., Erismann I.E., Ellis P.S., Bishop C.E., Agulnik A.I.,
RA Simpson E., Mitchell M.J.;
RT "Identification of a mouse male-specific transplacental antigen,
RT H-Y.";
RL Nature 376:695-698(1995).
-1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
-1- produced by alternative splicing.
-1- MISCELLANEOUS: Escapes X-inactivation.
-1- SIMILARITY: CONTAINS 1 ARID DOMAIN.
-1- SIMILARITY: CONTAINS 1 JmJc.
-1- SIMILARITY: CONTAINS 1 JMUN DOMAIN.
-1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
-1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF127245; AAD53049.1; -  
 DR EMBL: Z29651; CAA82759.1; -  
 DR EMBL: L29563; AAA62384.1; -  
 DR EMBL: AK008105; BAB25462.1; -  
 DR EMBL: AK011577; BAB27111.1; -  
 DR EMBL: AF039894; AAB96762.1; -  
 DR MGD: MG1:99781; Smcx.  
 DR InterPro: IPR001606; ARID.  
 DR InterPro: IPR003347; TF\_JmJc.  
 DR InterPro: IPR003349; TF\_JmJN.  
 DR InterPro: IPR004198; ZnF\_C5HC2.  
 DR InterPro: IPR001965; ZnF\_PHD.  
 DR Pfam: PF00628; PHD; 2.  
 DR Pfam: PF01388; ARID; 1.  
 DR Pfam: PF02373; JmJc; 1.  
 DR Pfam: PF02375; JmJN; 1.  
 DR Pfam: PF02928; zf-C5HC2; 1.  
 DR SMART: SM00501; BRIGHT; 1.  
 DR SMART: SM00545; JmJN; 1.  
 DR SMART: SM00249; PHD; 2.  
 DR PROSITE: PS01359; ZF\_PHD\_1; 2.  
 DR PROSITE: PS00016; ZF\_PHD\_2; 1.  
 DR ZINC-finger; Repeat; Alternative splicing.  
 FT DOMAIN 13 59 JmJN.  
 FT ZN\_FING 76 184 ARID.  
 FT DOMAIN 326 372 PHD-TYPE 1.  
 FT ZN\_FING 501 617 JmJc.  
 FT ZN\_FING 1187 1248 JmJc.  
 FT VARSPLIC 1364 1366 MISSING (IN ISOFORM 2) (BY SIMILARITY).  
 FT CONFICT 3 3 L -> M (IN REF. 5).  
 FT CONFICT 26 26 D -> N (IN REF. 4).  
 FT CONFICT 111 111 D -> N (IN REF. 4).  
 FT CONFICT 249 249 G -> A (IN REF. 4; BAB27111).  
 FT CONFICT 1068 1068 L -> P (IN REF. 2).  
 FT SEQUENCE 1554 AA; 175322 MW; A5E47BF5D1F5BAC CRC64;

Alignment Scores:  
 Pred. No.: 1 37 Length: 1554  
 Score: 57.00 Matches: 9  
 Percent Similarity: 64.00% Conservative: 7  
 Best Local Similarity: 36.00% Mismatches: 3  
 Query Match: 35.19% Indels: 6  
 Gaps: 1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x SMCX\_MOUSE (1-1554)

OY 80 AGTTGGTCAAGTGGATTGATTTATGCGCTGCCCATTCGCGATCCCATGATG 21  
 Db 1234 Alatrptpolturpaspfrlryspheleucys-----Prolaucys 1247  
 OY 20 ATTGCATCACAGG 6  
 Db 1248 MetAigSerArgArg 1252

RESULT 11  
 SMCX\_HUMAN  
 ID SMCX\_HUMAN STANDARD: PRT; 1560 AA.  
 AC P41229;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Smcx protein (Xel169 protein).  
 GN SMCX OR XE169.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=94214434; PubMed=8162017;  
 RA Wu J., Ellison J., Salido E., Yen P., Mohandas T., Shapiro L.J.;  
 RT "Isolation and characterization of XE169, a novel human gene that  
 RT escapes X-inactivation".  
 RL Hum. Mol. Genet. 3:153-160(1994).  
 RN [2]  
 RP SEQUENCE OF 280-344 FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=95038739; PubMed=7951230;  
 RA Aguinik A.I., Mitchell M.J., Martel M.-G., Ayner P.A.,  
 RA Lerner J.L., Bishop C.E.;  
 RT "A novel X gene with a widely transcribed Y-linked homologue escapes  
 RT X-inactivation in mouse and human".  
 RL Hum. Mol. Genet. 3:879-884(1994).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Expressed in all tissues examined. Highest  
 CC levels found in skeletal muscle.  
 CC -1- MISCELLANEOUS: Escapes X-inactivation.  
 CC -1- SIMILARITY: CONTAINS 1 ARID DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 JmJc DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 JmJN DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.  
 CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L25270; AAA61302.1; -  
 DR EMBL: Z29650; CAA82758.1; -  
 DR Genew; HGNC:11114; SMCX.  
 DR MIM: 314690; -  
 DR InterPro: IPR001606; ARID.  
 DR InterPro: IPR003347; TF\_JmJc.  
 DR InterPro: IPR003349; TF\_JmJN.  
 DR InterPro: IPR004198; ZnF\_C5HC2.  
 DR InterPro: IPR001965; ZnF\_PHD.  
 DR Pfam: PF00628; PHD; 2.  
 DR Pfam: PF01388; ARID; 1.  
 DR Pfam: PF02373; JmJc; 1.  
 DR Pfam: PF02375; JmJN; 1.  
 DR Pfam: PF02928; zf-C5HC2; 1.  
 DR SMART: SM00501; BRIGHT; 1.  
 DR SMART: SM00545; JmJN; 1.  
 DR SMART: SM00249; PHD; 2.  
 DR PROSITE: PS01359; ZF\_PHD\_1; 2.  
 DR PROSITE: PS00016; ZF\_PHD\_2; 1.  
 DR ZINC-finger; Repeat; Alternative splicing.  
 FT DOMAIN 13 59 JmJN.  
 FT ZN\_FING 76 184 ARID.  
 FT ZN\_FING 326 372 PHD-TYPE 1.  
 FT ZN\_FING 501 617 JmJc.  
 FT ZN\_FING 1187 1248 JmJc.  
 FT VARSPLIC 1370 1372 MISSING (IN ISOFORM 2).  
 FT CONFICT 342 342 C -> Y (IN REF. 2).  
 FT SEQUENCE 1560 AA; 175804 MW; 78913F010DBCCB1 CRC64;

Alignment Scores:  
 Pred. No.: 1 37 Length: 1560  
 Score: 57.00 Matches: 9  
 Percent Similarity: 64.00% Conservative: 7  
 Best Local Similarity: 36.00% Mismatches: 3  
 Query Match: 35.19% Indels: 6  
 Gaps: 1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x SMCX\_HUMAN (1-1560)

OY 80 AGTGGTCAAGTGGATTTGATTTATGCGCGTCCGCAATTCGCGATCCCGCATGATG 21  
 Db 1234 AATirpTgIurTpasPThLysPheUcys-----ProUcys 1247  
 OY 20 ATTCGATCAGAGG 6  
 Db 1248 MetArgSerArgArg 1252

## RESULT 12

T2M1\_METJA STANDARD; PRT; 222 AA.  
 ID T2M1\_METJA  
 AC 058391;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE TYPE II restriction enzyme MjaI (EC 3.1.21.4) (Endonuclease MjaI)  
 DE (R.MjaI).  
 GN MJAIR OR MJ0984.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 OC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 OX [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=6888087;  
 RA But C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sult G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kervaege A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reisch C.I.,  
 RA Overbeek R., Kirkness E.F., Weisslock K.G., Merrick J.M., Nguyen D.,  
 RA Scott J.L., Geoghegan N.S.M., Meldrum J.F., Fuhrmann J.L., Nguyen D.,  
 RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Bodrovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii.";  
 RL Science 273:1058-1073(1996).  
 RN [2]

## CHARACTERIZATION

RP Noren C.J., Roberts R.J., Patti J., Byrd D.R., Morgan R.D.;  
 RT "Method for screening restriction endonucleases.";  
 RL Patent number WO9911821, 11-MAR-1999.  
 CC -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE CTAG  
 AND CLEAVES AFTER C-1 (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give  
 CC specific double-stranded fragments with terminal 5'-phosphates.  
 CC -1- SIMILARITY: STRONG, TO M.THERMOFORMICUM MTHZIR.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: U67541; AAB98987.1;  
 DR REBASE: 1221; MjaI.  
 DR TIGR: MJ0984;  
 KW Hydrolyase; Endonuclease; Nuclease; Restriction system;  
 KW Complete proteome.  
 SQ SEQUENCE 222 AA: 25962 MW: 9A254A208AB8AC02 CRC64;

## Alignment Scores:

Pred. No.: 3.52 Length: 222  
 Score: 54.00 Matches: 12  
 Percent Similarity: 72.73% Conservative: 4  
 Best Local Similarity: 54.55% Mismatches: 6  
 Query Match: 32.93% Indels: 0  
 DB: 1 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x T2M1\_METJA (1-222)

OY 19 ATCATGATGGGATCGGATGGGACGACGACATTAATCAATCCACTGACGAA 78  
 Db 166 lIeIleIleasPaspLysGluIleProValSerIleLysSerLysfRnTyValasMn 185  
 OY 79 CTGAG 84  
 Db 186 GluLys 187

## RESULT 13

ERR2\_HUMAN STANDARD; PRT; 500 AA.  
 ID ERR2\_HUMAN  
 AC 095718; O9HCB4;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta)  
 DE (ERR-beta) (Estrogen receptor-like 2) (ERR beta-2).  
 GN ESRRB OR NR3B2 OR ESRRB2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 OX [1]

RP SEQUENCE FROM N.A.  
 RC Tissue-Testis;  
 RX MEDLINE=99173874; PubMed=10072763;  
 RA Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,  
 RA Liu Q., Caskey C.T.;  
 RT Identification of two hERR2-related novel nuclear receptors utilizing  
 RT bioinformatics and inverse PCR.";  
 RL Gene 228:101-109(1999).  
 RN [2]

RP SEQUENCE OF 264-500 FROM N.A.  
 RA Rowen L., Madan A., Qin S., Baradaran L., Birditt B., Bloom S.,  
 RA Dows M., Dickhoff R., Fleetwood P., Harrison G., Kaur A., Madan A.,  
 RA Nesbitt R., Traicoff R., Hood L.;  
 RT "Sequencing of human chromosome 14q24.3 region.";  
 RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 CC NR3 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF094517; AAC99409.1;  
 DR EMBL: AC008050; AAG17472.1;  
 DR HSSP: P03372; IHCO.  
 DR TRANSFAC: T02766;  
 DR GeneW: HGNC:3473; ESRRB.  
 DR MIM: 602167;  
 DR InterPro: IPR000536; Hormone\_rec\_11g.  
 DR InterPro: IPR001723; Steroid\_receptor.  
 DR InterPro: IPR001628; znfCsteroid.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; zf-C4\_1.  
 DR PRINTS: PR00398; STRDHOMONER.  
 DR PRINTS: PR00047; STROIDINGER.  
 DR ProDom: PD000035; znfCsteroid; 1.  
 DR SMART: SM00430; HOL1; 1.  
 DR SMART: SM00399; znfC4; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;  
 KW Zinc-finger.  
 KW DNA\_BIND 103 168 NUCLEAR RECEPTOR-TYPE.  
 FT ZN\_FING 103 123 C4-TYPE.  
 FT ZN\_FING 139 163 C4-TYPE.

SQL SEQUENCE 500 AA; 55619 MW; C219C84D914DF1C6 CRC64;

# Alignment Scores:

Score: 3.72 Length: 500  
Percent Similarity: 54.00 Matches: 14  
Best Local Similarity: 57.14% Conservative: 2  
Query Match: 50.00% Mismatches: 6  
DB: 32.93% Indels: 6  
Gaps: 2

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ERR2\_HUMAN (1-500)

OY 1 TCGACCTGCTGATGATGCA-----ATCATCATGGGATGCGATGCGACGCGAC 51

DB 244 CysAspLeuAlaAspArgGluLeuValIleIleGly-----TTPalaTylsHis 260

OY 52 ATAAATCCAAATCCCACTTGAAC 75

DB 261 IleProGlyPheSerSerLeuSer 268

## RESULT 14

RADC\_BACHD STANDARD; PRT; 232 AA.

AC 09K8H4; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

GN RADC OR BH3032.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_TaxID=86665;

RN [1]

RC STRAIN=C-125 / JCM 9153;

RA MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horiuchi K.,

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus

halodurans and genomic sequence comparison with Bacillus subtilis."

CC Nucleic Acids Res. 28:4317-4331(2000).

CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RADC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>

CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).

CC EMBL: AP001517; BAB06751.1; -

DR EMBL: AP001517; BAB06751.1; -

DR InterPro: IPR000445; Hhm.

DR InterPro: IPR001405; RADC.

DR Pfam: PF00633; Hhm; 1.

DR ProDom: PD007415; RADC; 1.

DR TIGRFAMs: TIGR00608; RADC; 1.

DR PROSITE: PS01302; RADC; 1.

KW DNA repair; Complete proteome.

SQL SEQUENCE 232 AA; 26403 MW; 4542153B5C6010D2 CRC64;

Alignment Scores:

Score: 4.22 Length: 232

Percent Similarity: 53.50 Matches: 11

Best Local Similarity: 69.57% Conservative: 5

Query Match: 47.83% Mismatches: 6

DB: 32.62% Indels: 1

Gaps: 1

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x RADC\_BACHD (1-232)

OY 4 GACCTGCTGATGATGATCATTCATTTGGGATGCGATGCGACGCGACATTAATCCAA 63

DB 208 GluLeuLeuAspHisValIleIleGlyAspArgTrpPheValSer---LeuYsgLulys 226

OY 64 TCCCACTTG 72

DB 227 GlyHisLeu 229

## RESULT 15

ERR2\_RAT STANDARD; PRT; 433 AA.

AC P11475; 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Steroid hormone receptor ERR2 (Estrogen-related receptor, beta) (ERR-

beta) (Estrogen receptor-like 2).

GN ERRB OR NR3B2 OR ESR2 OR ERR2.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RC TISSUE=Heart;

RA MEDLINE=8812546; PubMed=3267207;

RA Giguere V., Yang N., Segal P., Evans R.M.;

RT "Identification of a new class of steroid hormone receptors."

RL Nature 331:91-94(1988).

CC [2]

CC SHOWS THAT REF.1 ORIGINATES FROM RAT.

CC MEDLINE=99173874; PubMed=10072763;

CC Chen F., Zhang Q., McDonald T., Davidoff M.J., Bailey W., Bai C.,

CC Liu Q., Caskey C.T.;

CC "Identification of two hERR2-related novel nuclear receptors utilizing

CC bioinformatics and inverse PCR."

CC Gene 228:101-109(1999).

CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Nuclear (potential).

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

CC -1- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO ORIGINATE FROM HUMAN

CC BUT WAS LATER SHOWN (REF.2) TO BE DERIVED FROM RAT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>

CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).

CC EMBL: X51417; CA93779.1; -

DR EMBL: X51417; CA93779.1; -

DR HSPD: P03372; HCO.

DR InterPro: IPR000536; Hormone\_rec.119.

DR InterPro: IPR001723; Sterhorm\_receptor.

DR InterPro: IPR001628; Znf\_C4steroid.

DR Pfam: PF00104; hormone\_rec; 1.

DR Pfam: PF00105; zf-C4; 1.

DR PRINTS: PR00398; STRDHORMER.

DR PRINTS: PR00047; STROIDFINGER.

DR ProDom: PD000035; Znf\_C4steroid; 1.

DR SMART: SM00430; HOL1; 1.

DR SMART: SM00399; Znf\_C4; 1.

DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.

KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;

KW Zinc-finger.

FT DNA\_BIND 103 168 NUCLEAR RECEPTOR-TYPE.

FT ZN\_FING 103 123 C4-TYPE.

FT ZN\_FING 139 163 C4-TYPE.

SQL SEQUENCE 433 AA; 48287 MW; EE2C4C5B2P9A3E13 CRC64;

## Alignment Scores:

Pred. No.:	5.26	Length:	433
Score:	53.00	Matches:	11
Percent Similarity:	58.33%	Conservative:	3
Best Local Similarity:	45.83%	Mismatches:	10
Query Match:	32.32%	Indels:	0
DB:	1	Gaps:	0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ERR2\_RAT (1-433)

QY	1	TGTGACCTCTGTGATCGATCATCATTTGGGGATCGCGAATGCGCAGCCGACATATAAATCC	60
Db	244	CysaspLeuAlaaspArgGluLeuValPheLeuIleSerTrpAlaLysHisIleProGly	263
QY	61	AAATCCCACTTG	72
Db	264	PheSerAsnLeu	267

Search completed: April 21, 2003, 18:49:20  
Job time : 4.07696 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 16:30:20 ; Search time 3.16753 Seconds

(without alignments)  
7572.189 Million cell updates/sec

Title: us-09-513-151-3\_copy\_1121\_1210

Perfect score: 164

Sequence: 1 TGTGACCTCTGTGATGCAAT.....TGAACCACTGAGCAAGAAAGA 90

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+np.model -DEV=xlh  
-O=/cgn2\_1/USPRO.spool/US09513151/runat.15042003.141143.26358/app.query.fasta\_1.2446  
-DB=A.GeneSeq.101002 -QFMT=fastan -SUFFIX=np2p.rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151 -ECGN.1.1.84 -runat.15042003.141143.26358 -NCPU=3  
-NO.XLPHY -NO.MAP -LARGEOUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A.GeneSeq.101002:\*

```

1: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneSeq/geneSeqp-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	411	23	Novel human protei
2	64	39.0	302	21	Human ORFX ORF1724
3	63	38.4	617	23	Human ovarian anti
4	63	38.4	864	22	Novel human diagno
5	61	37.2	314	22	Drosophila melanog
6	60	36.6	60	23	Human ORFX protein
7	60	36.6	430	20	Amino acid sequenc
8	58	35.8	72	21	Arabidopsis thalia
9	58	35.8	72	21	Arabidopsis thalia
10	58	35.8	121	21	Arabidopsis thalia
11	58	35.4	578	22	Arabidopsis thalia
12	58	35.4	578	22	Drosophila melanog
13	58	35.4	884	22	Drosophila melanog
14	57.5	35.1	935	22	Drosophila melanog
15	57	34.8	737	21	Breast and ovarian
16	57	34.8	765	22	Human transcriptio
17	57	34.8	786	22	Human polypeptide
18	57	34.8	842	22	Human polypeptide
19	57	34.8	898	22	Human polypeptide
20	57	35.2	1539	18	Human SMC2 protein
21	56	34.1	269	22	Human protein sequ
22	56	34.1	346	22	Drosophila melanog
23	55.5	33.8	539	22	Drosophila melanog
24	55	33.5	652	22	Drosophila melanog
25	55	33.5	789	22	Drosophila melanog
26	54	32.9	123	21	Human secreted pro
27	54	32.9	187	21	Human secreted pro
28	54	32.9	433	21	Human secreted pro
29	54	32.9	500	20	Human nuclear rece
30	54	32.9	500	20	Human nuclear rece
31	54	32.9	508	21	Human nuclear rece
32	54	32.9	508	20	Human nuclear rece
33	53.5	32.6	238	22	Human novel secret
34	53.5	32.6	1424	22	Human polypeptide
35	53.5	32.6	1464	22	Human polypeptide
36	53	32.3	340	22	Human polypeptide
37	53	32.3	341	22	Human polypeptide
38	53	32.3	418	21	Human nuclear rece
39	53	32.3	418	21	Human nuclear rece
40	53	32.3	433	9	Sequence of human
41	53	32.3	433	21	Human ERbeta prot
42	53	32.3	435	19	Steroid hormone re
43	53	32.3	435	21	Human oestrogen re
44	53	32.3	458	20	Human nuclear rece
45	53	32.3	458	21	Human ERgamma pro

## ALIGNMENTS

RESULT 1  
ABB97295 standard; Protein: 411 AA.

27-JUN-2002 (first entry)

Novel human protein SEQ ID NO: 563.

Human; antihaemetic; vulnary; antiinflammatory; immunomodulator;  
antifertility; cerebroprotective; cytoskeletal; rheumatic; gene therapy;  
neuroprotective; antiparkinsonian; protein therapy; EST;  
expressed sequence tag.

Homo sapiens.

W0200222660-A2.

PD 21-MAR-2002.  
 XX 10-SEP-2001; 2001MO-US26015.  
 XX 11-SEP-2000; 2000US-0659671.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YF, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;  
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;  
 DR WPI: 2002-292408/33.  
 DR N-PSDB: ABN32481.  
 XX  
 PT An isolated polynucleotide for treating diseases associated with its  
 PT encoded polypeptide such as cancer and multiple sclerosis -  
 XX  
 PS Example 2; SEQ ID NO 563; 509pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences of 444  
 CC novel human proteins. These were isolated from expressed sequences tags  
 CC (ESTs). They can be used to stimulate cell growth, to regulate  
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth  
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat  
 CC multiple sclerosis, to regulate activin or inhibin e.g. to treat  
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat  
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions  
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.  
 CC Parkinson's disease. The present sequence is a protein of the invention.  
 XX  
 SQ Sequence 411 AA:  
 Alignment Scores:  
 Pred. No.: 3,56e-18 Length: 411  
 Score: 164.00 Matches: 30  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ABB9795 (1-411)  
 QY 1 TGTGACCTGTGATGGAATCATCATTTGGGATCGGAAATGGGACGACATAAATCC 60  
 DB 341 CysAspLeucCysAspArgIleIleGlyAspArgGlnTrrAlaIleIleSer 360  
 QY 61 AATCCACTTGACCACTGAAGAAGA 90  
 DB 361 LysSerHisLeuAsnGlnLeuLysArg 370  
 RESULT 2  
 AAB41960  
 ID AAB41960 standard; Protein: 302 AA.  
 XX  
 AC AAB41960;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human ORFX ORF1724 polypeptide sequence SEQ ID NO:3448.  
 XX  
 KW Human: open reading frame; ORFX: detection; cytosstatic; hepatotropic;  
 KW vulnary: antiporatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antinflammatory;  
 KW antiviral; antibacterial; antifungal; antineumatic; antihypoid;  
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;

KW thrombosis; contraceptive.  
 XX  
 OS Homo sapiens.  
 XX WO200058473-A2.  
 XX  
 XX 05-OCT-2000.  
 XX  
 PF 31-MAR-2000; 2000MO-US08621.  
 XX  
 PF 31-MAR-1999; 99US-0127607.  
 XX  
 PR 02-APR-1999; 99US-0127636.  
 PR 05-APR-1999; 99US-0127728.  
 PR 30-MAR-2000; 2000US-0540763.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI: 2000-602362/57.  
 DR N-PSDB: AAC76169.  
 XX  
 PT Novel nucleic acids and peptides derived from open reading frame X,  
 PT useful for treating e.g. cancers, proliferative disorders,  
 PT neurodegenerative disorders and cardiovascular disease -  
 XX  
 PS Claim 11; Page 2624; 5507pp; English.  
 XX  
 CC AAC74446 to AAC7606 encode the proteins given in AAB40237 to AAB43397,  
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX  
 CC sequences have activities such as: cytosstatic; hepatotropic; vulnary;  
 CC antiporatic; antiparkinsonian; nootropic; neuroprotective;  
 CC osteopathic; anticonvulsant; antiarthritis; immunosuppressant;  
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
 CC antinflammatory; antibacterial; antiviral; antifungal; antineumatic;  
 CC antihypoid; and antianaemic. The sequences can be used for determining  
 CC the presence of or predisposition to, or preventing or treating  
 CC pathological conditions associated with an ORFX-associated disorder. The  
 CC nucleic acids can be used to express ORFX proteins in gene therapy  
 CC vectors. The proteins and nucleic acids may be used to treat cancers,  
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
 CC nocturnal haemoglobinuria, antinflammatory disease; to enhance  
 CC coagulation; to inhibit thrombosis; and as a contraceptive.  
 XX  
 SQ Sequence 302 AA:  
 Alignment Scores:  
 Pred. No.: 0.12 Length: 302  
 Score: 64.00 Matches: 10  
 Percent Similarity: 65.52% Conservative: 9  
 Best Local Similarity: 34.48% Mismatches: 10  
 Query Match: 39.02% Indels: 0  
 DB: 21 Gaps: 0  
 US-09-513-151-3\_COPY\_1121\_1210 (1-90) x AAB41960 (1-302)  
 QY 1 TGTGACCTGTGATGGAATCATCATTTGGGATCGGAAATGGGACGACATAAATCC 60  
 DB 83 CysLysValCysCysAlaLeuLeuIleSerGlnSerGlnLysLeuAlaIleTyrglnSer 102  
 QY 61 AATCCACTTGACCACTGAAGAAGA 87  
 DB 103 LysLysHisAlaAsnLysValLysArg 111  
 RESULT 3  
 ABB41976  
 ID ABB41976 standard; Protein: 617 AA.

XX AC ABP41976;  
 XX XX 22-AUG-2002 (first entry)  
 DT XX  
 XX XX Human ovarian antigen HAO63, SEQ ID NO:3108.  
 DE XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antinflammatory; gynaecological; reproductive.  
 KW XX  
 OS Homo sapiens.  
 XX OS  
 XX PN WO200200677-A1.  
 XX PD 03-JAN-2002.  
 XX XX 07-JUN-2001; 2001WO-US18569.  
 XX PF 07-JUN-2001; 2000US-209467P.  
 XX PR 07-JUN-2000; 2000US-209467P.  
 XX XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Birse CE, Rosen CA;  
 XX PI  
 XX DR WPI: 2002-147878/19.  
 XX DR N-PSDB: ABQ55053.  
 XX XX  
 XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer), immune disorders, cardiovascular disorders and  
 PT neurological diseases -  
 XX PT  
 XX PS Claim 11; SEQ ID NO 3108; 2922pp; English.  
 XX XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antibodies against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the  
 CC identification of individuals and in forensic analysis, and the  
 CC polypeptides may be used as food additives or to prepare antibodies  
 CC useful in disease diagnosis, drug targeting and phenotyping. The present  
 CC sequence represents a human ovarian antigen of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX CC  
 SO Sequence 617 AA;

Alignment Scores:  
 Pred. No.: 0.203  
 Score: 63.00  
 Percent Similarity: 60.87%  
 Best Local Similarity: 39.13%  
 Query Match: 38.41%  
 DB: 23  
 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ABP41976 (1-617)

QY 1 TGTGACCTGTGATCGAATCATTCATGGGATCGCGAATGGACGACATTAATCC 60  
 ||| ::|||  
 DB 69 CysSerIleCysAspLeuProValHisSerAsnLysGluTrpSerClnHisIleAsnGly 88  
 ||| |||||  
 QY 61 AATCCAC 69  
 |||||  
 DB 89 AlasSerHis 91

RESULT 4  
 ABG18701  
 ID ABG18701 standard: Protein; 864 AA.  
 XX AC ABG18701;  
 XX XX 18-FEB-2002 (first entry)  
 XX DT  
 XX DE Novel human diagnostic protein #18692.  
 XX XX  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX OS  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US08631.  
 XX PR 31-MAR-2000; 2000US-0540217.  
 XX PR 23-AUG-2000; 2000US-0649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX PI  
 XX DR WPI: 2001-639362/73.  
 XX DR N-PSDB: AAS82888.  
 XX XX  
 XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX PT  
 XX PS Claim 20; SEQ ID NO 49060; 103pp; English.  
 XX XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations in  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and



CC specification, but was obtained in electronic format directly from WIPO

CC sequences can be used in gene therapy. ORF sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,  
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,  
CC osteoarthritis, neurodegenerative disorders, disorders related to organ  
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic  
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester  
CC storage disease, various immune deficiencies and disorders, infectious  
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid  
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host  
CC disease and autoimmune inflammatory eye disease. ORF proteins are also  
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,  
CC bone degenerative disorders, or periodontal disease, and for gut  
CC protection or regeneration and treatment of lung or liver fibrosis,  
CC reperfusion injury in various tissues and conditions resulting from  
CC systemic cytokine damage.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 60 AA;

Alignment Scores:  
Pred. No.: 0 402 Length: 60  
Score: 60.00 Matches: 11  
Percent Similarity: 58.33% Conservative: 3  
Best Local Similarity: 45.83% Mismatches: 10  
Query Match: 36.59% Indels: 0  
DB: 23 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ABP04585 (1-60)

QY 18 AATCATCATTTGGGATCGGAGATGGCGACACATAAATCCAAATCCACTTGAACCA 77  
||| ||||| ::||| ||||| :: ||||| |||||  
Db 24 AAsAAGtIsrPrProthraArgAlaGlySerProthraArgProleuPrThGluPro 43

QY 78 ACAGAGAGAAAG 89  
||| |||||  
Db 44 GlyProGluLys 47

RESULT 7

ID AAY02526 standard; Protein: 430 AA.

AC AAY02526;

DT 15-JUL-1999 (first entry)

XX Amino acid sequence of the GRO-1 protein.

XX gro-1 operon; gro-1 gene; gop-1 gene; gop-2 gene; gop-2 gene;

KW hap-1 gene; cancer; aging; longevity; tumour formation;

KW physiological clock.

XX Caenorhabditis elegans.

XX WO910482-A1.

PD 04-MAR-1999.

PF 20-AUG-1998; 98MO-CA00803.

PR 25-AUG-1997; 97CA-2210251.

PA (UYMC-) UNIV MCGILL.

DR WPI; 1999-190615/16.

DR N-PSDB; AAX36072.

XX Barnes T, Hekimi S, Lakowski B, Lemieux J;

PT Molecular identity of the gro-1 gene - useful for cancer diagnosis

PT and/or prognosis, and where compounds affecting encoded proteins are

PT useful for enhancing longevity of a host and inhibiting tumour

PT formation

XX Cialm 11; Page 47-48; 93pp; English.  
PS The present sequence represents the GRO-1 protein. The specification  
XX describes the five genes of the Caenorhabditis elegans gro-1 operon  
CC (AAX36071). The operon contains the gro-1 gene (AAX36072), the  
CC gro-1 gene (AAX36074), the gop-2 gene (AAX36075), the gop-2 gene  
CC (AAX36075), and the hap-1 gene (AAX36077). The gro-1 gene can be used in  
CC a method for the diagnosis and/or prognosis of cancer in a patient.  
CC Transgenic mice containing a gene knock-out of a murine gene homologue  
CC of the gro-1 gene are useful as models of aging and cancer. The  
CC proteins encoded by the genes are useful for identifying compounds  
CC that affect the enzymatic activity of these proteins, in order to  
CC enhance longevity of a host and inhibit tumour formation. The gro-1  
CC gene, together with the gop-1, gop-2, gop-3 and hap-1 genes enables  
CC study of a physiological clock.  
XX  
SQ Sequence 430 AA;

Alignment Scores:  
Pred. No.: 0 593 Length: 430  
Score: 60.00 Matches: 9  
Percent Similarity: 53.33% Conservative: 7  
Best Local Similarity: 30.00% Mismatches: 14  
Query Match: 36.59% Indels: 0  
DB: 20 Gaps: 0

US-09-513-151-3\_COPY\_1121\_1210 (1-90) x AAY02526 (1-430)

QY 1 TGTGACCTCTGTGATCGAATCATTTGGGATCGGAGCGACGACATAAATCC 60  
||| ||||| ::||| ||||| :: ||||| |||||  
Db 394 CysGluIleCysAsnIleSerMetThrGlyLysAspAsnTrpGlnLysIleAspGly 413

QY 61 AATGCCACTTGACGACCACTGAAGAAAGA 90  
||| |||||  
Db 414 LysLysHisLysHisHisAlaLysGlnLys 423

RESULT 8

ID AAG56792 standard; Protein: 72 AA.

AC AAG56792;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 73087.

XX Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-01322407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134270.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142390.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142820.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143524.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149829.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158322.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
  
Alignment Scores:  
Pred. No.: 0.893 Length: 72  
Score: 58.00 Matches: 12  
Percent Similarity: 51.61% Conservative: 4  
Best Local Similarity: 38.71% Mismatches: 9  
Query Match: 35.80% Indels: 6  
Gaps: 1  
DB: 21  
  
US-09-513-151-3\_COPY\_1121\_1210 (1-90) x AAG56792 (1-72)  
  
OY 80 AGTGGTCACAGTGGGATTGATTTATGTCGCCCTCCCATTCGCATCCCATGATG 21  
DB 27 SetrtpargleutrgpvalaIProphetrcysvalhshsapglnleProleugly 46  
OY 20 ATT-----CGATCAGAGG 6  
DB 47 llelleleuleutleargglunrgerdlnarg 57  
  
RESULT 9  
AAG60375  
ID AAG60375 standard; Protein; 7? AA.  
XX AAG60375;  
AC  
XX  
XX 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 78196.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX EP1033405-A2.  
PN  
XX  
XX  
PD 06-SEP-2000.  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142970.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.

```

PR 20-JUL-1999; 990S-0144632.
PR 20-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 21-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 26-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 23-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158363.
PR 13-OCT-1999; 990S-0158393.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.

```

```

PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Alignment Scores:
Pred. No.: 0.893
Score: 58.00
Percent Similarity: 51.61%
Best Local Similarity: 38.71%
Query Match: 35.80%
DB: 21

US-09-513-151-3_COPY_1121_1210 (1-90) x AAG60375 (1-72)

OY 80 AGTGGTCAAGCGCATTTGATGTGCGCCATCGGACATCCCATGANG 21
    ||||| |||:||||| ||| ||| |||:|||||
Db 27 SetTTPArgLeuTpcIuValProPheTpcYsValHisHisApsInIleProLeugly 46
OY 20 ATT-----CGATCAGAGG 6
    ||| ||||| ||||| ||||| ||||| |||||
Db 47 IleIleGluLeuIleArgGluArgSerGlnArg 57

RESULT 10
AAG56791
ID AAG56791 standard; Protein; 121 AA.
XX
AC AAG56791;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 73086.
XX
KW protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.

```

PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148665.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154039.  
PR 16-SEP-1999; 99US-0154779.  
PR 20-SEP-1999; 99US-0155139.  
PR 22-SEP-1999; 99US-0155486.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.

PR 06-APR-1999; 99US-0128234.

PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144331
PR	16-JUL-1999;	9905-0144085
PR	16-JUL-1999;	9905-0144084
PR	16-JUL-1999;	9905-0144083
PR	15-JUL-1999;	9905-0144062
PR	15-JUL-1999;	9905-0144061
PR	15-JUL-1999;	9905-0143542
PR	13-JUL-1999;	9905-0143541
PR	12-JUL-1999;	9905-0142920
PR	12-JUL-1999;	9905-0142919
PR	08-JUL-1999;	9905-0142903
PR	06-JUL-1999;	9905-0142390
PR	02-JUL-1999;	9905-0142055
PR	01-JUL-1999;	9905-0142154
PR	01-JUL-1999;	9905-0141842
PR	30-JUN-1999;	9905-0141287
PR	29-JUN-1999;	9905-0140991
PR	28-JUN-1999;	9905-0140823
PR	24-JUN-1999;	9905-0140354
PR	23-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0138859
PR	21-JUN-1999;	9905-0138176
PR	18-JUN-1999;	9905-0138175
PR	18-JUN-1999;	9905-0138750
PR	18-JUN-1999;	9905-0138453
PR	18-JUN-1999;	9905-0138452
PR	18-JUN-1999;	9905-0138451
PR	18-JUN-1999;	9905-0138450
PR	18-JUN-1999;	9905-0138449
PR	18-JUN-1999;	9905-0138448
PR	18-JUN-1999;	9905-0138447
PR	18-JUN-1999;	9905-0138446
PR	18-JUN-1999;	9905-0138445
PR	16-JUN-1999;	9905-0138457
PR	16-JUN-1999;	9905-0138456
PR	16-JUN-1999;	9905-0138455
PR	16-JUN-1999;	9905-0138454
PR	16-JUN-1999;	9905-0138453
PR	16-JUN-1999;	9905-0138452
PR	16-JUN-1999;	9905-0138451
PR	16-JUN-1999;	9905-0138450
PR	16-JUN-1999;	9905-0138449
PR	16-JUN-1999;	9905-0138448
PR	16-JUN-1999;	9905-0138447
PR	16-JUN-1999;	9905-0138446
PR	16-JUN-1999;	9905-0138445
PR	16-JUN-1999;	9905-0138444
PR	16-JUN-1999;	9905-0138443
PR	16-JUN-1999;	9905-0138442
PR	16-JUN-1999;	9905-0138441
PR	16-JUN-1999;	9905-0138440
PR	16-JUN-1999;	9905-0138439
PR	16-JUN-1999;	9905-0138438
PR	16-JUN-1999;	9905-0138437
PR	16-JUN-1999;	9905-0138436
PR	16-JUN-1999;	9905-0138435
PR	16-JUN-1999;	9905-0138434
PR	16-JUN-1999;	9905-0138433
PR	16-JUN-1999;	9905-0138432
PR	16-JUN-1999;	9905-0138431
PR	16-JUN-1999;	9905-0138430
PR	16-JUN-1999;	9905-0138429
PR	16-JUN-1999;	9905-0138428
PR	16-JUN-1999;	9905-0138427
PR	16-JUN-1999;	9905-0138426
PR	16-JUN-1999;	9905-0138425
PR	16-JUN-1999;	9905-0138424
PR	16-JUN-1999;	9905-0138423
PR	16-JUN-1999;	9905-0138422
PR	16-JUN-1999;	9905-0138421
PR	16-JUN-1999;	9905-0138420
PR	16-JUN-1999;	9905-0138419
PR	16-JUN-1999;	9905-0138418
PR	16-JUN-1999;	9905-0138417
PR	16-JUN-1999;	9905-0138416
PR	16-JUN-1999;	9905-0138415
PR	16-JUN-1999;	9905-0138414
PR	16-JUN-1999;	9905-0138413
PR	16-JUN-1999;	9905-0138412
PR	16-JUN-1999;	9905-0138411
PR	16-JUN-1999;	9905-0138410
PR	16-JUN-1999;	9905-0138409
PR	16-JUN-1999;	9905-0138408
PR	16-JUN-1999;	9905-0138407
PR	16-JUN-1999;	9905-0138406
PR	16-JUN-1999;	9905-0138405
PR	16-JUN-1999;	9905-0138404
PR	16-JUN-1999;	9905-0138403
PR	16-JUN-1999;	9905-0138402
PR	16-JUN-1999;	9905-0138401
PR	16-JUN-1999;	9905-0138400
PR	16-JUN-1999;	9905-0138399
PR	16-JUN-1999;	9905-0138398
PR	16-JUN-1999;	9905-0138397
PR	16-JUN-1999;	9905-0138396
PR	16-JUN-1999;	9905-0138395
PR	16-JUN-1999;	9905-0138394
PR	16-JUN-1999;	9905-0138393
PR	16-JUN-1999;	9905-013839

```

PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157553.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.

```

```

PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 0.99
Score: 58.00
Percent Similarity: 51.61%
Best Local Similarity: 38.71%
Query Match: 35.80%
DB: 21

US-09-513-151-3_COPY_1121_1210 (1-90) x AAG60374 (1-121)

QY 80 AGTGTGATCACTGAGATTGATTTATGTCGCTGCCCATTCGCAATCCCAATGATG 21
    ||||| |||:|:| ||| ||| |||:|:|
DB 76 SerrTPargLeutrpGluValProphetpCysValHshIsAspGlnIleProLeugly 95
    ||| |-----CGATCACAGAG 6
QY 20 ART-----
DB 96 IleIleGluLeuIleArgGluArgSerGlnArg 106

RESULT 12
ABB58393
ID ABB58393 standard; Protein; 578 AA.
XX
AC ABB58393;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 1971.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
    pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX

```



DR WPI: 2001-656860/75.  
DR N-PSDB; ABL02496.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure: SEQ ID NO 1971; 21pp + Sequence Listing: English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (AB101840-AB116175).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 578 AA:  
  
Alignment Scores:  
Pred. No.: 1.35 Length: 578  
Score: 58.00 Matches: 11  
Percent Similarity: 62.50% Conservative: 4  
Best Local Similarity: 45.83% Mismatches: 7  
Query Match: 35.37% Indels: 2  
DB: 22 Gaps: 1  
  
US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ABB58393 (1-578)  
QY 1 TGTGACCTCTGTGATGCATCATTCGCGAATGGCGACGACATAAATCC 60  
DB 315 CysGluLeuCysGlyArgMetPheSerSerArgAspGluTrpSerIleHis-----Ala 332  
OY 61 AAATCCCACTTG 72  
DB 333 LysSerHisLeu 336  
  
RESULT 13  
ABB61358  
ID ABB61358 standard; Protein; 884 AA.  
XX  
AC ABB61358;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 10866.  
XX  
KM Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI: 2001-656860/75.  
DR N-PSDB; ABL05461.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure: SEQ ID NO 10866; 21pp + Sequence Listing: English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins  
CC (AB101840-AB116175).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 884 AA:  
  
Alignment Scores:  
Pred. No.: 1.47 Length: 884  
Score: 58.00 Matches: 10  
Percent Similarity: 53.57% Conservative: 5  
Best Local Similarity: 35.71% Mismatches: 13  
Query Match: 35.37% Indels: 0  
DB: 22 Gaps: 0  
  
US-09-513-151-3\_COPY\_1121\_1210 (1-90) x ABB61358 (1-884)  
QY 1 TGTGACCTCTGTGATGCATCATTCGCGAATGGCGACGACATAAATCC 60  
DB 255 CysGluLeuCysAspValThrCysThrGlyThrAspAlaValAlaValAlaValGly 274  
OY 61 AAATCCCACTTGAAACCACTGAAG 84  
DB 275 AlaLysHisGlnIleValValLys 282  
  
RESULT 14  
ABB58283  
ID ABB58283 standard; Protein; 935 AA.  
XX  
AC ABB58283;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 1641.  
XX  
KM Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI: 2001-656860/75.  
DR N-PSDB; ABL02386.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure: SEQ ID NO 1641; 21pp + Sequence Listing: English.

	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL1840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
Sequence	935 AA:
 Alignment Scores:  	
Pred. No.:	1.79 Length: 935
Score:	57.50 Matches: 13
Percent Similarity:	53.85% Conservative: 1
Best Local Similarity:	50.00%
Mismatches:	11
Gaps:	1
Query Match:	35.068 Indels: 1
Differences:	22 Gaps: 1
OY	US-09-513-151-3_COPY_1121_1210 (1-90) x ABB58283 (1-935)
Db	1 TGTNACCTCGT---GATGCATCATCATTGGGAGTCGCCGATGGCACGCATAA 57
Oy	
DB	783 CysasprrcysleusprilelialiserAspsSergltirPalalysargllelys 802
OY	58 TCCAATCCCATTTGAAC 75
Db	:::
DB	803 LeugluysPheargasn 808
RESULT 15	
AAB59023	ID AAB59023 standard; Protein; 737 AA.
XX	AAB59023;
XX	
DE	27-MAR-2001 (first entry)
XX	Breast and ovarian cancer associated antigen protein sequence SEQ ID 731
KW	Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive; neurotrophic; neuropsychotropic; antiviral; antileptocytic; hepatotrophic; antimetabolic; antiinflammatory; anticancer; vulnereary; anticoagulant; antibacterial; antifungal; antiparasitic; cardiac; immune disorder; Addison's disease; allergy; autoimmune hemolytic anaemia; autoimmune thyroiditis; diabetes mellitus; Crohn's disease; multiple sclerosis; rheumatoid arthritis; ulcerative colitis; cardiovascular disorder; wound healing; neurological disease.
XO	Homo sapiens.
PN	WO200055173-A1.
PD	21-SEP-2000.
PF	08-MAR-2000; 2000MO-US05881.
PR	12-MAR-1999; 99US-0124270.
PA	(HUMA-) HUMAN GENOME SCI INC.
PI	Rosen CA, Ruben SM;
DR	WIPI; 2000-611515/58. N-PADB; AAF21926.
PT	New human breast and ovarian cancer associated gene sequences and the polypeptides encoded by these genes, useful in the prevention, treatment and diagnosis of cancer, Immune disorders, cardiovascular disorders and neurological diseases -

```

XX  Claim 11: Page 1193-1196; 1299pp; English.
XX
XX  Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
XX  proteins AAB58711 - AAB59128. The DNA and protein sequences are
XX  associated with breast and ovarian cancer. Included in the invention
XX  sequences AAF22032 - AAF22040 and AAB59129 which are used in the
XX  isolation and characterisation of the DNA and protein sequences of t
XX  invention. The breast and ovarian cancer associated DNA, protein, ag
XX  or antagonist sequences exhibit cytostatic; immunosuppressive;
XX  neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX  antidiabetic; antiinflammatory; antitumor; vulnerable; anticonvulsant
XX  antibacterial; antifungal; antiparasitic and cardiant activity. The
XX  polynucleotide and protein sequences are used in the diagnosis of ca
XX  particularly breast and ovarian cancer. The nucleic acid sequences,
XX  proteins, agonists and agonists may also be used in the diagnosis,
XX  prevention and treatment of immune disorders e.g. Addison's disease,
XX  allergies, autoimmune hemolytic anemia, autoimmune thyroiditis,
XX  diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
XX  arthritis and ulcerative colitis; cardiovascular diseases such as
XX  myocardial ischaemia; wound healing; neurological diseases such as
XX  cerebral anoxia and epilepsy; and infectious diseases.
XX
SQ  Sequence      737 AA:

Alignment Scores:
Pred. No.:      2.07      Length:      737
Score:          57.00      Matches:      11
Percent Similarity: 53.57%      Conservative:  4
Best Local Similarity: 39.29%      Mismatches:  13
Query Match:      34.76%      Indels:      0
DB:              21      Gaps:        0

US-09-513-151-3_COPY_1121_1210 (1-90) x AAB59023 (1-737)
QY      1  TGTGACCTCTGTGATGGAATCATCTATTGGGATCGCGAATGGCGAGCATATAAATCC 60
      |||:::|||||||  |||  ::|||
Db      47  CysgluLeucysAspValSerCysThrGlyAlaaspAlaGlyAlaHisIleArgGly 66
QY      61  AAATCCCACTTGAAACCACTGAAG 84
      |||  ::|||
Db      67  AlalysHisGlnLysValValLys 74

Search completed: April 21, 2003, 18:48:18
Job time : 5.16753 secs

```



```

Db 21 ArgThrLeuProLeuValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40
QY 131 TTGAGCTAGAGCCGCGGCTCGGAGATGCTACGGCTGACTCCATGCAAGTGTAT 190
Db 41 LeuGlnLeuGlyGlnArgLeuGlyGlyIleValSerIleAspSerMetGlnValTyr 60
QY 191 GAAGGCTTAGACATCATCACCACCAAGGTTTCGCCCAAGAGAGAAATCTGCCGCAC 250
Db 61 GlnGlyLeuAspIleIleThrAsnLysValSerIleGlnGlnValGlyIleCysArgHis 80
QY 251 CACATGATCAGCTTTGTGATCCCTGTGTGACCAATATACAGTGTGAGCTTCAGAAAT 310
Db 81 HisMetIleSerPheValAspProLeuValThrAsnThrThrValValAspPheArgAsn 100
QY 311 AGAGCAACTGCTGTGATTGAA-----GATATATTGCCCGA 346
Db 101 ArgAlaThrAlaLeuIleSerLeuGlyLysAlaAlaIleValLysPheAspIlePheAlaArg 120
QY 347 GACAAATTCCTATTGTTGTGGAGAGAAACAATTATACATTTGAAATCTGCTCGGAAA 406
Db 121 AspLysIleProIleValValGlyGlyThrAsnTyrTrlIleGlnSerLeuLeuTrpLys 140
QY 407 GTTCTGTCAATACCAAGCCCAAGAGATGGGACTGAGAAAGTATGATGACCCGAAAGTG 466
Db 141 ValLeuValAsnThrLysProGlnGlnMetGlyThrGlyLysValIleAspArgLysVal 160
QY 467 GAGTTTAAAAAGAGATGGTCTTGTACTTCACAAAGCCTTAAGCCAGGTGGAGCCGAA 526
Db 161 GlnLeuGlnLysGlnAspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGln 180
QY 527 ATGGCTGCAGAGTGCATCCACATGACAAAGCAAGTGGCCAGAGACTTGCAGAGTTT 586
Db 181 MetAlaAlaLysLeuHisProHisAspLysArgLysValAlaArgSerLeuGlnValPhe 200
QY 587 GAAGAAACAGGAATCTCTCATAGTAATTTCTCCATGCTACATACAGGAAGAGTGTGT 646
Db 201 GlnGlnThrGlyLysSerHisSerGlnPheLeuHisArgGlnHisThrGlnGlnGly 220
QY 647 GGTCCCTTTGGAGGCTCTCTGAAAGTGTCTTAACCTTGCATCCTTTGGCTTCATGCTGAC 706
Db 221 GlyProLeuGlyGlyProLeuLysPheSerAsnProCysIleLeuTrpLeuHisAlaAsp 240
QY 707 CAGGCACTTAGATGAGCCCTTGATAGAGAGGCTGATACATGCTTGGCTGGGCTC 766
Db 241 GlnAlaValLeuAspGlnValArgLeuAspLysArgValAlaAspMetLeuAlaIleGlyLeu 260
QY 767 TTGGAGAACTAGAGATTTTTCACAGACGCTATATATAGAGAAATGTTTGGAAATATAGC 826
Db 261 LeuGlnGlnLeuArgAspPheHisArgArgLysGlnGlnLysAsnValSerGlnLysAsnSer 280
QY 827 CAGGACTATACATGATGATCTTCATCCATCAATTTGGCTTCAAGGAATTTACAGAGTACTG 886
Db 281 GlnAspLysGlnHisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGlnTyrLeu 300
QY 887 ATCACTAGAGGAAATGACACTGAGAGACTAGTATACAGCTTCTTAAGAAAGCA----- 940
Db 301 IleThrGlnGlnLysCysThrIleGlnIleThrSerAsnGlnLeuLeuLysGlyIleGln 320
QY 940 ----- 940
Db 321 AlaLeuLysGlnValThrLysArgTyrAlaArgLysGlnAsnArgTrpValLysAsnArg 340
QY 941 -----CCTGTCCATTTGTCGCCCTGCTATAGGCTTAGAGATCTGATGATGC 988
Db 341 PheLeuSerArgProGlyProIleValAlaProValIleGlyIleGlnGlnValSerAspVal 360
QY 989 TCGAAGTGGAGAGAGTCTGTTCTTGAACCTGCTTGAATTCGTCAAAATTTTCATCAG 1048
Db 361 SerLysTrpGlnGlnSerValLeuGlnProAlaLeuGlnIleValAlaGlnSerPheIleGln 380
QY 1049 GGGCACAAGCCTTACAGCCACTCCCAATAAAGATGCCATACATGAGCTGAGAACACAGAGA 1108
Db 381 GlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnIleValAsnLysArg 400

```

```

QY 1109 AGTTATACCTGTGTGACTCTGTGATCGAATCATCATTTGGGATGCCAATGGCAGCG 1168
Db 401 SerTyrHisLeuLysAspLysPheCysAspArgIleIleIleGlyAspArgGlnTrpAlaAla 1228
QY 1169 CACATAAATTCACAAATCCACTTGAACCACTGAGAAAGAAAGAGATGGACTAGAT 1228
Db 421 HisIleLysSerLysSerHisLeuAsnGlnLeuLysLysArgValGlnLeuAspSerAsp 440
QY 1229 GCTGTCAACACATAGAACTCAGAGTGTTCGCCAGACTATACAAAGAACCTAAAGGG 1288
Db 441 AlaValAsnThrIleGlnSerGlnSerValSerProAspHisAsnLysGlnProLysGln 460
QY 1289 AAGGATCCCGAGGAGAGATGATCAAGACTGAATGACGCGTT 1333
Db 461 LysGlySerProGlyGlnAsnAspGlnGlnLeuLysCysSerVal 475

RESULT 2
US-60-453-050-7968
; Sequence 7968, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7968
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-7968

Alignment Scores:
Pred. No.: 1,02e-172 Length: 475
Score: 2234.00 Matches: 438
Percent Similarity: 92.42% Conservative: 1
Best Local Similarity: 92.21% Mismatches: 2
Query Match: 62,49% Indels: 34
DB: 7 Gaps: 2

US-09-513-151-3 (1-2041) x US-60-453-050-7968 (1-475)
QY 11 ATGGCGTCCGTGGCGGCTGCAGCAGAGTCTCTGTGGCAGTGGGCTCAGAGGCTGCAAA 70
Db 1 MetAlaSerValAlaAlaAlaArgAlaValProValIleGlySerGlyLeuArgGlyLeuGln 20
QY 71 CGGACCTACTCTTGTAGTGAATTTCTGGGGCCACGGGCAACCGGCAATTCACAGCTGGCG 130
Db 21 ArgThrLeuProLeuValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40
QY 131 TTCCAGCTAGAGCCAGCGGCTCGGGGCTGAGATGCTGACGCTGACTCATGAGAGTGTAT 190
Db 41 LeuGlnLeuGlyGlnArgLeuGlyGlyIleValSerIleAspSerMetGlnValTyr 60
QY 191 GAAGGCTTAGACATCATCACCACCAAGGTTTTCGCCCAAGAGAGAAATCTGCCGCAC 250
Db 61 GlnGlyLeuAspIleIleThrAsnLysValSerIleGlnGlnValGlyIleCysArgHis 80
QY 251 CACATGATCAGCTTTGTGATCCCTGTGTGACCAATATACAGTGTGAGCTTCAGAAAT 310
Db 81 HisMetIleSerPheValAspProLeuValThrAsnThrThrValValAspPheArgAsn 100
QY 311 AGAGCAACTGCTGTGATTGAA-----GATATATTGCCCGA 346
Db 101 ArgAlaThrAlaLeuIleSerLeuGlyLysAlaAlaIleValLysPheAspIlePheAlaArg 120
QY 347 GACAAATTCCTATTGTTGTGGAGAGAAACAATTATACATTTGAAATCTGCTCGGAAA 406
Db 121 AspLysIleProIleValValGlyGlyThrAsnTyrTrlIleGlnSerLeuLeuTrpLys 140

```

QY	407	GTCTTGTCAATACCAAGCCCCAGAGATGGGCACTGAGAAATGATTTGACCGAAATGTG	466
Db	141	ValLeuValAsnThrLysProGlnGluMetGlyThrGlyValIleAspArgLysVal	160
QY	467	GAGCTTAAAGAGAGATGGTCTGTACTTCAACAAAGCGCTTAAGCGAGGTGACCCAA	526
Db	161	GluLeuGlyLysGlnAspGlyLeuValIleuHisLysArgLeuSerGlnIleAspProGlu	180
QY	527	ATGGCTCCCAAGCTGCATCCATGACAAACCAAGTGGCCAGAGAGCTTGCAGATTTT	586
Db	181	MetIleAlaLysLeuHisProHisAspLysArgGlyValAlaArgSerLeuGlnIleAlaPhe	200
QY	587	GAGAAACACAGAAATCTCTACTACTGAATTTCTCCATGCTGCACATACAGAGAAAGTGT	646
Db	201	GluIleuThrAlaGlyIleSerHisSerGlnPheLeuHisArgGlnHisThrGluIleuGly	220
QY	647	GGTCCCTTGAGAGTCCCTGGAAGTCTGTAAACCCCTGCATCCCTTGGCTCATGCTGAC	706
Db	221	GlyProLeuGlyGlyProLeuLysPheSerAsnProCysIleLeuThrPheuHisIleAsp	240
QY	707	CAGGCACTTATAGATGAGCGCTTGATTAAGAGGTGGATGACATGCTTCTGTGCGCTC	766
Db	241	GlnIleValLeuAspGlnArgLeuAspLysArgValAspAspMetLeuAlaIleu	260
QY	767	TTTGAGAGAACTAAGACATTTTTCACACACCGCTATATACAAACAATGTTTCGAAAAATGC	826
Db	261	LeuIleuIleuLeuArgAspPheHisAspArgLysArgGlnLysAsnValSerIleuAsnSer	280
QY	827	CAGACATATACATGATGGTATCTCCCAATCAATATGGCTTCAAGAAATTCACAGATACCTG	886
Db	281	GlnAspArgGlnHisGlyIlePheGlnSerIleGlyPheLysGlnPheHisGluArgLeu	300
QY	887	ATCACTGAGGAAAAATGCACACTGGAGACTAGTAAACCAAGCTTCTAAAGAAAGA-----	940
Db	301	IleThrGlnGlyLysCysThrLeuGlnThrSerAsnGlnLeuLeuLysGlyIleGlu	320
QY	940	-----	940
Db	321	AlaLeuLysGlnValThrLysArgLysAlaArgLysGlnAsnArgTrpValLysAsnArg	340
QY	941	-----CCGTGCTCCATTTGTCCCCCTGCTGTATAGCTTATAGAGTATCTGATGTC	988
Db	341	PheLeuSerArgProGlyProIleValProProValTyrGlyLeuGlnValSerAspVal	360
QY	989	TTCGAATGGAGAGAGTGTGTTCTTGAACCTGCTTGAATGTGTCAAAGTTTCATCCAG	1048
Db	361	SerLysTrpGlnGlnSerValLeuGlnProAlaLeuGlnIleValGlnSerPheIleGln	380
QY	1049	GGCCACAAGCTCACAGCCACTCCCAATAAAGATCCCTATACAAAGAAAGCTGAGAAACAAGA	1108
Db	381	GlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnIleAsnLysArg	400
QY	1109	AGTTATCACTGTGTGACCTGTGTGATCGAATCATCATTTGGGATGCGCAATGGGCAAGC	1168
Db	401	SerTyrHisLeuLysAspLeuLysAspArgGlyIleIleGlyLysArgGlyIleTrpIleAla	420
QY	1169	CACATTAATATCCAAATCCCACTGAAACCAATGAAAGAAAGAAAGAAAGTGGACATCAAT	1228
Db	421	HisIleLysSerLysSerHisLysLeuAsnGlnLeuLysLysArgArgTrpLeuAspSerAsp	440
QY	1229	GCTTCAACACACATAGAAAGTACAGAGTGTTTCCCGACATATACAAAGAAAGACCTAAAGG	1288
Db	441	AlaValAsnThrIleGlnSerGlnSerValSerProAspHisAsnLysGlnIleProLysGln	460
QY	1289	AAGGATCCCCAGAGGACAGAAATGATCAAGACGTGAATGCAAGCGTT	1333
Db	461	LysGlySerProGlyGlnAsnAspGlnGlnLeuLysCysSerVal	475

[illegible]

QY	899	AAACACACCTGGAGACAGTACGACGCTTCAAGCAAGGA-----	940
Db	241	lyscysthrleugluthrserasnleuileuylsglyleclualaleuylsgln	260
QY	940	-----	940
Db	261	valthrlysarqgtyrlyalaarglysglnasnaagttprvallyasnaarphleuSERarg	280
QY	941	CCTGGTCCCATGTGTCCCCCTGTCTATGGCTTAGAGATTCGATGTCGAAGTGGGAG	1000
Db	281	ProglyProilevalProprovaltyrglyleuglivalserasphalserlystrgln	300
QY	1001	GAGTGTGTTCTTGAACCTGCTCTTGCAATCGTGCAGAAATTTCATCCAGGGCCACAGCCT	1060
Db	301	gluservalleugluproalaaleuglunillevalglnserphelleclnglyhislvspro	320
QY	1061	ACAGGCACCTCCCAATAAGTGGCATACATGAAGCTGACACAGAGAAAGTTATCCACCTG	1120
Db	321	ThrAlathnProileuysheprOtyraangluAlaGlnaslnysargserlyrhlslseu	340
QY	1121	TGTGACCTCTGTGATCGAATCATCATTTGGGGATCCGGCAATGGGACGCACATATAAATCC	1180
Db	341	Cysasphleucysaspharglellellelleglyaspharglutrphalaiahnshlelysser	360
QY	1181	AAATCCCACTTGTAACCACTGAAGAAGAAAGAGAGATTGGACTCAGATGCTGTCAACACC	1240
Db	361	lysSerhIshleuasnGlnleuylslysarqargargleuaspserasphalavalanThr	380
QY	1241	ATAGAAAGTCAGAGTGTTCCTCCAGACTATAACAAAGAACCTRAAGGAAAGGATCCCA	1300
Db	381	IlegluserGlnserValserProasphIsaanlysglnuprolyselublysglyserpro	400
QY	1301	GGGCAATGATCAAGAGCTGAATGCAGCGTT 1333	
Db	401	glyglinasnaspGlnleuileuylsCysSerVal 411	

```

RESULT 4
US-60-453-135-7967
; Sequence 7967, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7967
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-7967

Alignment Scores:
Pred. No.:      2,1e-72      Length:      221
Score:          997.00      Matches:      193
Percent Similarity: 87.78%      Conservative: 1
Best Local Similarity: 87.33%      Mismatches:  1
Query Match:     27.89%      Indels:      26
DB:              7          Gaps:         1
US-09-453-151-3 (1-2041) x US-60-453-135-7967 (1-221)

```

Oy	749	ATGCTGCCTGGCGGGCTCTTGGAGGACCTAAAGATTTCACAGACGCTATATACGAG	808
Db	1	Mcttcaataataacgcytneuteneutugtneuaatgaasphenisatrgatrgyrannglnlvs	20
Oy	809	AATGTTTCGGAAATATGCCAGACGACTATCAACATGGTATTTCCATCAATGTCCTCAG	868
Db	21	Asnvalsercgluasnerfinspryrglnhslgyllephenelnerllleglypnelvs	40

QY	869	GAATTTCACGACAGTACTGATCAGTACTGAGGGAAATATGACACCTGAGACACTAGTATACCAGCTT	928
Db	41	GluthehisglutylrLeuilethmglnghlyscysthrLeuglutrhSerAsnclndeu	60
QY	929	CTAAGAAAGCA-----	940
Db	61	LeulysylselylleglualaleuylsglnValthrLysArgtyrAlaArgLysclnAsn	80
QY	941	-----CCTGTCCCATTTGTCCCCCTGTCTATGGC	970
Db	81	ArgtrpValylsAsnargPheleuSerArgproGlyProilevalProProvaltyrGly	100
QY	971	TTAAGTATCTATGCTCTGAGTGGGAGAGCTGTGCTTGAACCTGCTTGAATC	1030
Db	101	LeuclvalseraspalserlystrpGlnghlserValLeuglnprAlaLeuglnlle	120
QY	1031	GTCGAAGTTTATCCAGGGCCACAAGCCTACAGCCACTCCAAATAAGTCCATACAT	1090
Db	121	ValclnserhelleinglnghlyshlystrProthrAlaThrProilelsmecProtrlysn	140
QY	1091	GAACCTGAGAACAGAGACAGTATTCACCTGTGTGACCTCTGTGATGCAATCATCTATTGGG	1150
Db	141	GlnlagaLslnslnylsargserlyrHisLeuylsAspLeuylsAspArgllellellegly	160
QY	1151	GATGCCAAATGGGCGACGGCGACATAAATATCCCAATCCCACTTATACCACTGAAAGAAAGA	1210
Db	161	AspArgglutrrpAlaalaahisileLysSerLysSerHisLeuAsnclndeuLyslystrg	180
QY	1211	AGAAGATTTGACATCAGATGCTGTCAACACCATPAGAAAGTCAGAGTGTTCGCCACACAT	1270
Db	181	ArgtrgLeuaspseraspAlaValanThrillegluSerGlnSerValserProaspHis	200
QY	1271	AACAAAGAACCTAAAGGGAAGGAGATCCCGACGGCGAGCAATGATCAAGAGCTGAATGCAGC	1330
Db	201	AsnlysglnProlysglnLysglySerProglyGlnAsnAspGlnclndeuLyscysSer	220
QY	1331	GTT 1333	
Db	221	Val 221	

```

RESULT 5
US-60-453-050-7967
: Sequence 7967, Application US/60453050
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: SENOSIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001457
: CURRENT APPLICATION NUMBER: US/60/453,050
: CURRENT FILING DATE: 2003-03-10
: NUMBER OF SEQ. ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7967
: LENGTH: 221
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-60-453-050-7967

```

Alignment Scores:		
Pred. No.:	2,1e-72	Length: 221
Score:	997.00	Matches: 193
Percent Similarity:	87.78%	Conservative: 1
Best local Similarity:	87.33%	Mismatches: 1
Query Match:	27.89%	Indels: 26
DB:	7	Gaps: 1
US-09-513-151-3 (1-2041) x US-60-453-050-7967 (1-221)		
QY	749	ATGTTTCTGCTGGGCTCTTGSGAGAACTAAGAGATTTCACAGACGCTATTATCAGAG 808
db	1	MetleuulalaaclgyleuLeuGluGluLeuAlaGAsPheNisAAlGAlGylAsnGlnIlys 20

```

Y 809 AATGTTTCGAAATATAGCCAGACATATCAATCAATGCTATCTTCCATTCATTAATGGCTCAAG 868
      |||
      |||
      |||
Db 21 AsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePheGlnSerIleGlyPheLys 40
      |||
      |||
      |||
Y 869 GAATTCACGAGTACCTGATCACTAGGAGAAATACACACTGGAGACTAGTAAACAGCTT 928
      |||
      |||
      |||
Db 41 GluPheHisGluTyrLeuIleIleThrGlnGlySerCysThrLeuGlnThrSerAsnGlnLeu 60
      |||
      |||
      |||
Y 929 CTTAAGAAAGGA----- 940
      |||
      |||
      |||
Y 61 LeuLysLysGlyIleGluAlaLeuLysGlnValThrLysArgTyrAlaArgLysGlnAsn 80
      |||
      |||
      |||
Y 941 -----CCTGGTCCCATTTGCCCCCTGTCTATGGC 970
      |||
      |||
      |||
Db 81 ArgTyrValLysAsnArgPheLeuSerArgProGlyProIleValProProValTyrGly 100
      |||
      |||
      |||
Y 971 TTAAAGGATATCTGATGCTCTCGAAGTGGAGAGAGTCTGTCTTGAAACCTGCTTGAAATC 1030
      |||
      |||
      |||
Db 101 LeuGluValSerAspValSerLysTyrGlnGlnSerValLeuGlnProAlaLeuGlnIle 120
      |||
      |||
      |||
Y 1031 GTGCAAACTTCATCATCCAGGCGCACAAAGCCTACAGCCACTCCATATAAGATGCCATCAAT 1090
      |||
      |||
      |||
Db 121 ValGlnSerPheIleGlnGlyHisLysProThrAlaThrProIleLysMetProTyrAsn 140
      |||
      |||
      |||
Y 1091 GAAGCTGGAACACAGAGAAGTTATCATCCCTGTGTGACCTCTGTGATCGAATCATCATTTGGG 1150
      |||
      |||
      |||
Db 141 GluAlaGlnAsnLysArgSerTyrHisLysCysAspLeuLysAspArgIleIleGly 160
      |||
      |||
      |||
Y 1151 GATGCCGAATGGGCGACGCCACATTAATCCAAATCCCACTTGAAACCACTGAAGAAAGA 1210
      |||
      |||
      |||
Db 161 AspArgGluTyrPheAlaAlaHisIleLysSerLysSerHisLeuAsnGlnLeuLysLysArg 180
      |||
      |||
      |||
Y 1211 AGAAGATTGGACTCGATGCTGTCTCAACACCATAGAAGAGCAGAGTGTTCCTCCAGACTAT 1270
      |||
      |||
      |||
Db 181 ArgAlaGluAspSerAspAlaValAsnThrIleGlnSerGlnSerValSerProAspHis 200
      |||
      |||
      |||
Y 1271 AACAAAGAAGACCTAAAGGAGGATGCCAGGCGCAGAGAATGATCAAGAGTGAATCAGC 1330
      |||
      |||
      |||
Db 201 AsnLysGlnProLysGlnLysGlnSerProGlyGlnAsnAspGlnLeuLysCysSer 220
      |||
      |||
      |||
Y 1331 GTT 1333
      |||
      |||
      |||
Db 221 Val 221

RESULT 6
US-60-453-135-7969
: Sequence 7969, Application US/60453135
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele
: APPLICANT: IAKOUBOVA, Olga.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: FILE REFERENCE: CLO01456
: CURRENT APPLICATION NUMBER: US/60/453,135
: NUMBER OF SEQ ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7969
: LENGTH: 85
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-453-135-7969

```

Alignment Scores:	
Pred. No.:	1.58e-27
Score:	443.00
Percent Similarity:	98.83%
Best Local Similarity:	97.65%
Query Match:	12.39%
DB:	
	7
	Gaps:
	0

QY	1079	ATGCCATACATTAACCTGGAACAAGACAAGTATACCTGTGACCTGTGGACGA	1138
Db	1	MePrCtGyrtsnGlnlaGlnaLnLysArgSerTYrHisLeuCyAspLeuCyAspArg	20
QY	1139	ATCATCATTTGGGATGGCCGATATGGGACGGCACATTAATTCAAATCCCATTTGACCA	1198
Db	21	lIeIleIleIlysprArgGtutrrpAlaAlaHisIleLysSerLysSerHisLeuasnIn	40
QY	1199	CTGAAGAAAGAAGAAGATTGAGCATGCTGCTCAACACCATAGAAAGTCAGAGTGT	1258
Db	41	LeuLysLysArgtrgHrghLeuAspSerAspAlaValAsnThrIleGlnSerGlnSerVal	60
QY	1259	TCCCGACACTATATACAAAGAACTAAAGGAGAGGATCCCCAGGCGACAGATGATCAAGAG	1318
Db	61	SerProAspHisAsnLysGlnProLysGlnLysGlySerProGlnGlnAsnAspGlnIn	80
QY	1319	CTGAATATGACGCGTT	1333
Db	81	LeuLysCysSerVal	85

```

RESULT 7
US-60-453-050-7969
: Sequence 7969, Application US/60453050
:
: GENERAL INFORMATION:
:
: APPLICANT: LUKE, May
:
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: STENOISIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001457
: CURRENT APPLICATION NUMBER: US/60/453,050
: CURRENT FILING DATE: 2003-03-10
: NUMBER OF SEQ. ID NOS: 82762
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 7969
:
: LENGTH: 85
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
US-60-453-050-7969

```

Alignment Scores:	
Pred. No.:	1,586-27
Score:	443.00
Percent Similarity:	98.82%
Best Local Similarity:	97.65%
Query Match:	12.33%
DB:	7
	Gaps: 0

```

OY 1079 ATCCATATACATGAAGCTGGAACCAAGAGAAAGTTATCACCTGGTGTACCTTCGTATCGA 1138
Db 1 MetProtyrAnGluAlaAGluAsnLysValGSetTyrHisLeucysAspLeucysAspArg 20
OY 1139 ATCATCATTTGGGCGATCCGCGAATGGGCGCCACACATATAAAATCCAAATCCACCTTGAACCA 1198
Db 21 IleIleIleGlyAspArgGluTrrPAlaAlaHisIleLysSerLysSerHisLeuAsnGln 40
OY 1199 CTGAAGAAAAAAGAAGATTGGACTCGATCGATCGTGTATACACCATATAAGCTAGAGTGT 1258
Db 41 LeuLysLysAspArgAlaGlyLeuAspSerAspAlaValAsnThrIleLysSerLysSerVal 60
OY 1259 TCCCCAGACATTTAAACAAGAACCTAAAGGAGAGGATCCCCAGGCGCAGATGATCAAGG 1318
Db 61 SerProAspHisAsnLysGluProLysGluLysGlySerProGlyLysAsnAspGlnGlu 80
OY 1319 CTGAATGCGACGCTT 1333
Db 81 LeuLysCysSerVal 85

RESULT 8
PCT-US02-36123-4332
; Sequence 4332, Application PC/TUS0236123
; GENERAL INFORMATION:

```

```

; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; APPLICANT: Russell, David P., and Zagursky, Robert J.
; TITLE OF INVENTION: Alloococcus ocellidis Open Reading Frames (ORFs) Encoding Polypep
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4332
; LENGTH: 315
; TYPE: PRT
; ORGANISM: Alloococcus ocellidis
PCT-US02-36123-4332

Alignment Scores:
Pred. No.: 2,29e-23 Length: 315
Score: 393.00 Matches: 103
Percent Similarity: 55.59% Conservative: 81
Best Local Similarity: 31.12% Mismatches: 113
Query Match: 10.99% Indels: 34
DB: 1 Gaps: 10

US-09-513-151-3 (1-2041) x PCT-US02-36123-4332 (1-315)
QY 83 CTTGATGATTCCTCGGGCCACGGGCGACCGCAATCCACCGCTGGCTTGACGCTAGGC 142
DB 5 llellevalllevalglyprothrghlyvalglylysthlylseuuserleuglualea 24
QY 143 CAGCGGCTCGGGCGGTGAGATCTGACGGCTGACCTGACGCTGATGAGGCTTGAAGCCCTAGAC 202
DB 25 Glnaspleuallglylgluilelleasnglyaspsermetglvaltyrghlyleuasp 44
QY 203 ATCATCCACAAGAGTTTCTGCCCAAGACAGAGAGATCTGCCGGCACCATGATGATCACC 262
DB 45 lleglythglylyvalthrghlyserghlylalylyleprohishlyseuasp 64
QY 263 TTGTGATCTCTGTGACCAATTCACAGAGGTGAGCTGACCTGCAATATGAGCAACTGCT 322
DB 65 lleleuglu---valglndlnuapryrservalserasphneginsertnglyatmet 83
QY 323 CTGATTAAGATATATTTGCCCGAGACAAATTCCTATTGTGTGGAGAAACCAATTAAT 382
DB 84 lysilleglalllevalalaserlysaasnargleuproillevalglylythrglyleu 103
QY 383 TACATTAATCTCTGCTCGGAAAGTCTT-----GTCAATACCAAGCCCGACGAGANG 436
DB 104 tyrlleghlyserleuiletyraspvalserhsglylylaspasplaglnproasptyrglu 123
QY 437 GGCACCTGAGAAAGTATGACCGCAAAAGTGAGCTTGAAGAGAGATGCTTGTACTT 496
DB 124 tyraalalalyseuaglnaspleu-----Alaspselngluylrglutyrtle 140
QY 497 CACAAAGCCCTAAGCCAGGTGAGCCGCAAAATGCTGACAGCTGCATCCACATGACAA 556
DB 141 Tprglnleuileasnharglyeusprrplysialaialalylllehsproksnshleu 160
QY 557 CGCAAGTGCCAGAGCTTGCAGTCTTTTGAAGAAACGAGATCTCTATGTAATTT 616
DB 161 valargythlilearglaleuagluvaltyrhlsvalthrglyluthrpheserperhe 180
QY 617 CTCATTCGCAACATAGAGAGGTGGTGGTCCCTTGAGAGTCCCTGAGATTTCTCT 676
DB 181 -----Glnaspslulysylslylpro-----AlaleuTYR 191
QY 677 AACCTTCATCTTGGCTTCATCTGACAGGACAGGATTCATGATGAGCGCTGGATAG 736
DB 192 GlnalaphellelleghlyeusaunthrasparserlyleuTYRaspargllesnthr 211
QY 737 AGGTGGATGACATCTTCTGCTGGGCTCTTGAGAGAACTAGAGATTTTCACAGACGC 796
DB 212 Argvalasplmetvalasplysglyleuagluval-----Lys 226

```

```

QY 797 TATATCAGAGAGATTTTCGGAATATACGACGATATCAATGATATCTTCATCA 856
DB 227 Tprleuallalalysthlyrghlyasplaglnserarg-----Arggly 241
QY 857 ATTGGCTTAAGAGATTTACAGATACGATCTGATCAGTGAAGAAATGCACTGAGACT 916
DB 242 lleghlytyrarglvalleuuserTYRleu-----Aspghlyaspselnthrphneglu 259
QY 917 AGTAACGACTCTTAAGAAAGAGAGCTGCTCCATGTCCCT-----GTCTATGCG 970
DB 260 AlalleargaspllelyslasnsnserarghlsTYRAlalyarglndlnuthrtrp 279
QY 971 TTAGAGTATCTGATCTCGAGTGAGGAGACTGTG-----CTTGA 1015
DB 280 Hlsasnarghthraspvalasplystrpyrtyraspleuileglinsprrglylndln 299
QY 1016 CCTGCTTGAATCTGCAAAAGTTTCATCCAG 1048
DB 300 Gluvalillellyaspyalllysaasnphleuaglu 310

RESULT 9
US-09-950-084-4328
; Sequence 4328, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: George S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/IC963052
; CURRENT APPLICATION NUMBER: US/09/950,084
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 4328
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-4328

Alignment Scores:
Pred. No.: 5.52e-19 Length: 319
Score: 339.00 Matches: 86
Percent Similarity: 55.64% Conservative: 62
Best Local Similarity: 32.33% Mismatches: 92
Query Match: 9.48% Indels: 26
DB: 5 Gaps: 7

US-09-513-151-3 (1-2041) x US-09-950-084-4328 (1-319)
QY 83 CTTGATGATTCCTCGGGCCACGGGCGACCGCAATCCACCGCTGGCTTGACGCTAGGC 142
DB 16 llellevalllevalglyprothrghlylaserghlylsthlylseuuserlleghlyleu 35

```



```

QY 143 CAGCGGCTGGCGGTGAGATCTGACGGCTGACGAGGTCTATGAGGCGCTAGAC 202
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36 LysArgIleAsnIleValIleIleSerGlyAspSerMetIleValIleTyrIleHisMetAsn 55
QY 203 ATCATCACAACAGGTTTCTGCGCAAGACAGAGAATCTGCGGCGACCATGATCAGC 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 IleGIYThrAlaLysValIleThrProGluGluMetAspIleLeuProIleHisIleLeuAsp 75
QY 263 TTGTGGATTCCTCTGTGACCAATTCACAGTGTGAGACTTCAGAAATATAGCAACGCT 322
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 IleLeuAsnProAspAspIlePheSerAlaTyrGluIleLeuAspGluAlaGluAsp 94
QY 323 CTGATTGAAGATATATTGGCCGAGACAAATTCCTATTGTGGGAGAGAACCAATTAT 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 95 LeuIleThrAspIleThrAsnAlaGlyValProIleIleAlaGlyIleGlyLeu 114
QY 383 TACATTGAATCTCTGCTGGAAGAGTTCTTCATATACCAAGCCCGAGAGATGGGCACT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TyrIleGlnSerLeuIleTyrAsnIleTyrGluIleGluAspIleThrValIleProAlaGln 134
QY 443 GAGAAAGTGATGACCGAAAGTGAG---CTTGAAGAGAGAGATGCTCTTACTTCAC 499
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 135 LeuSerIleValIleLysGlnLysLeuSerAlaLeuGluIleLeuAspAsnGlnIleLeuHis 154
QY 500 AAACGCGCTAAGCAGGTGAGCCGAGAAATGCTGCCAAGCTGCATCCACATGACAAAGC 559
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 AspTyrLeuAlaGlnPheAspAlaValSerAlaGluAsnIleHisProAsnAsnAlaGln 174
QY 560 AAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGA-----ATCTCATATGAGA 613
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ArgValLeuArgAlaIleGluTyrTyrLeuLysThrIleLysLeuSerAsnAlaGlys 194
QY 614 TTCTTCATGCTGACATACATACAGAGAAAGAGTGCTGCTCCCTGGAGCTCTCTAAGTTC 673
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 LysVal---GlnGlnPheThrGluAsn----- 202
QY 674 TCTAACCTTGATCCCTTGGCTTCATGCTGACAGGAGAGTCTAGATGAGCGCTTGAT 733
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 TyrAspThrLeuLeuLeuGlyIleGluMetSerIleGlyThrLeuTyrSerArgIleAsn 722
QY 734 AAGAGGTGGATGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 LysArgValAspIleMetLeuAspIleGlyLeuPheArgGluVal----- 237
QY 794 CGCTATATACAGAGATGTTTGGAGAAATACAGAGCATATCAACATGGATATCTTCAA 853
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 -----GlnGlnLeuValIleGluGlnGlyTyrGlnIleSerGlnSer-----MetGln 252
QY 854 TCAATTGGCTTCAAGAA 871
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 253 AlaIleGlyTyrIleGln 258

```

```

Score: 319.50 Matches: 88
Percent Similarity: 52.168 Conservative: 57
Best Local Similarity: 31.658 Mismatches: 92
Query Match: 8.948 Gaps: 41
DB: 6 Gaps: 11

US-09-513-151-3 (1-2041) x US-10-289-762-983 (1-342)
QY 83 CTGTAGTATCTTCGCGGCGCAGCGGCGGAGAAATCCAGCGGTGAGCTAGGC 142
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 34 IleValLeuLeuSerGlyProIleGlySerGlyThrAspValSerIleAlaLeuAla 53
QY 143 CAGCGGCTGGCGGTGAGATCTGACGGCTGACGAGGTCTATGAGGCGCTAGAC 202
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 PrometIleAspIleGluIleValSerValAspSerMetIleValIleGlnIleMetAsp 73
QY 203 ATCATCACAACAGGTTTCTGCGCAAGACAGAGAATTCGCGGCGACCATGATCAGC 262
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 IleGIYThrAlaLysValSerLeuLysAlaArgGlnIleLeuProIleHisIleLeuAsp 93
QY 263 TTGTGGATTCCTCTGTGACCAATTCACAGTGTGAGACTTCAGAAATATAGCAACGCT 322
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 ---IleArgHisValGlnGluProIleAsnValAlaAspPheTyrTyrGluAlaIleGln 112
QY 323 CTGATTGAAGATATATTGGCCGAGACAAATTCCTATTGTGGGAGAGAACCAATTAT 382
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 AlaCysGlnAsnIleLeuSerArgAsnIleValProIleLeuValIleGlyIleSerGlyPhe 132
QY 383 TACATTGAATCTCTGCTGGAAGAGTTCTTCATATACCAAGCCCGAGAGATGGGCACT 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 TyrPheHisAlaPheLeu-----SerIleProProIleGlyProAlaAla 147
QY 443 GAGAAAGTGATGACCGAAAGTGAG---CTTGAAGAGAGAGATGCTCT---GTACTT 496
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 AspProGlnIleArgGlnIleGluIleGluAlaIleLeuGluIleValIleSerAlaLeu 167
QY 497 CACAACGCGCTAAGCAGGTGAGCCGAGAAATGCTGCCAAGCTGCATCCACATGACAA 556
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 TyrGlnAspLeuLeuLeuLysAspProGluTyrAlaGlnThrIleThrIleAsnAlaPlys 187
QY 557 CGCAAGTGGCCAGAGCTTGCAAGTTTGAAGAAACAGA-----ATCTCATATG 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 AsnIleIleIleArgGlyLeuGlnIleIleGlnLeuThrGlyLysValSerAspHis 207
QY 611 GAATT-----CTCCATGCTCAACATACGAGAAAGGTGGTGGCCCTT 655
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 GluTyrAspIleValIleProLysAlaSerArgIleTyrGlyS----- 220
QY 656 GAGAGTCTCTGAAGTTCTTAACTTGGATCTTGG---CTTCATGCTGACCAAGCA 712
    ::::::::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 -----CysArgAlaThrPheLeuSerProGluThrGlu 231
QY 713 GTTCTATATGAGCGCTTGATGATGAGGTGATGATGCTGCTGCTGCGGCTTGGAG 772
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 232 PheLeuLysAsnAsnIleGlnMetArgCysGluIleMetLeuGlnIleGlyLeuGln 251
QY 773 GACTAAGAGATTTTCAACAGCGCTATATACAGAAAGATGTTTGGAAAAATAGCCAGAG 832
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 252 GluValAlaArgGlyLeu-----LeuAsnGln---GlyIleArgGluAsnPro----- 265
QY 833 TATCAACATGCTATCTCAATCAATGCTTCAAGAAATTCAGAGTACCTG 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 266 -----SerAlaPheLysAlaIleGlyTyrArgIleTyrIleGluPheLeu 280

```

Alignment Scores: 2.14e-17 Length: 342

```

: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4383
: LENGTH: 311
: TYPE: PRT
: ORGANISM: Enterococcus faecalis
US-09-134-000C-4383

Alignment Scores:
Pred. No.: 3,26e-15 Length: 311
Score: 292.50 Matches: 84
Percent Similarity: 50.68% Conservative: 65
Best Local Similarity: 28.57% Mismatches: 102
Query Match: 8.18% Indels: 43
DB: Gaps: 5 8

US-09-513-151-3 (1-2041) x US-09-134-000C-4383 (1-311)

QY 83 CTTGATGATGATCTCTGGGGCCAGGCGCAATCCAGCTGGCTGACGTAAGCC 142
   |||||
Db 4 VALLEVALILLEVALIGLYPROTHRALVALDIGLYUASNSERILEYSTRYCYSPHECY 23
   |||||
QY 143 CAGCGG---CTCGGGGGAGATGTCMGACGGCTGACTCCATGACAGTGTATGAAAGCCCA 199
   |||||
Db 24 GINLPSILILEMETERVALYSLEULESERGLYASPSERMETGLNVALYRPROLEUVAL 43
   |||||
QY 200 GACATCATCACCCAAC--AAGTTTCTCGCCCAAGAGCAGACAGATATGCCGACACATG 256
   |||||
Db 44 ARGLEUTPRPROARGLEULYSSETHRGLYTHRGILYLSGLINGLYPHEROHSITRYLEU 63
   |||||
QY 257 ATCAGCTTGTGATCCTCTTGTGACCAATTCACAGTGTGACTTCAGAAATAGAGCA 316
   |||||
Db 64 ILIAPSPCYSGNLINVALISERGLNTHR--TYRSERLAALAASPHEGLNLYSGLUVAL 82
   |||||
QY 317 ACTGCTGTGATGAGATATATTATTCGCCGAGCAAAATTCATATTGTTGGAGGAACC 376
   |||||
Db 83 ARGGLNLSILELYSGIULLEHTRGLYLSGLYLSLEUPROILELLEVALIGLYLYTHR 102
   |||||
QY 377 AATTATTCATGATGATCTCTGCTGGAAGTTCTTGTCAAATACCAAGCCCGAGAGATG 436
   |||||
Db 103 GLYLEUTYRILEGLINSERILEUZYRASPHE-----GINLEU 115
   |||||
QY 437 GGCACATGACAAAGTATGACCGA-----AAAGTGGAGCTTCAA 475
   |||||
Db 116 GLYSERARGGLNILEASPSERPROGLIULEARGLINTHRTYRASNLEUPHEALAGLU 135
   |||||
QY 476 AAGGAGGATGCTGTGCTACTTCACAAAGCCCTAAGCCAGTGGAGCCCAAAATGGCTGCC 535
   |||||
Db 136 GINLPSGLYASNGILALALEUTRPLEULEBUNGILNLYSAPROLEUALAALAASN 155
   |||||
QY 536 AAGCTGACATCCATGACCAAAAGCAAGTGGCGAGGAGCTTCAAGTTTGAAGAAACA 595
   |||||
Db 156 SEIRILEHLSPHASNASNARGLYSLYSVALILEARGALALEUGLIVALPHEASPLYSTR 175
   |||||
QY 596 GGAATCTCTCATATGGAATTTCTCCATGCTCAACATACGGAAGAAGTGTGTCCTCT 655
   |||||
Db 176 GITYRSERIELEUETHR-----PROLYS 183
   |||||
QY 656 GGAGGTCCCTCTGAAGTCTCTAACCCCTTCATCCTTGGCTCATGTCGACGAGCAGGT 715
   |||||
Db 184 GILNPSPROALAIRLEUTYRSPRYTRYTRILEULENGILYLEUGLINTHRSPARGALAEU 203
   |||||
QY 716 CTAGATGAGCGCTTGGATTAAGAGGTGATGATGATGCTGTGCTGGAGCTTTGGAGAA 775
   |||||
Db 204 LEUTYRGLIARGILEASNGILNARGVALASPGIMETMETTRHGLINGLYLEULEUGLU 223
   |||||
QY 776 CTAGACAGATTTTACACAGACGCTAATATCAAGAAGATGTTTCGAAATATGCCAGACTAT 835
   |||||
Db 224 ALATYSGINMETPHEGLINGINPROHLSIAGINLAALA----- 236
   |||||
QY 836 CAACATGATATTCCATGATGCTGGCTTCAGAGAAATTCACAGAGTACTATCACTGAG 895
   |||||
Db 237 -----GLINGLYILEGLYTRYLPSGLIUPHEPHEROTRYRPE-----SER 249

```

[illegible]

```

QY 656 GGAGGTCCTCGAGGTCCTCTAAACCCCTGGATCCTTGGCTTCATGTCGACGAGGAGT 715
Db 184 G1uylsPrOAlaIglLeuYlYrAspYlYrTlLeuLeuYlLeuU1uthrAspAlaLeu 203
QY 716 CTAGATGAGCCCTTGATTAAGAGGGTGAAGACATGCTTCTGCTGCGGCTTGAGGAAA 775
Db 204 LeuYlrg1uAgl1IleAsng1nArG1ValAspG1nMetethrG1uG1yLeuLeuG1uG1n 223
QY 776 CTAAGAGATTTTCACAGACGCTATATTCAGAGAGATGTTTCGGAAATAGCCAGACTAT 835
Db 224 A1a1ysG1nMetPheG1ng1nG1nPrOAlaG1nAla1a1a----- 236
QY 836 CAACATGGTATCTTCACATGATCGGCTTACAGAAATTCACAGACTGATCAGTAC 895
Db 237 -----G1ng1Y1IleG1Yr1YlG1uG1nUthPheProTYrPhe-----Ser 249
QY 896 GGAANAATGCACACTGAGACTAGTAAACAGCTTCTTAAGAAA 937
Db 250 G1yG1uG1nSerLeuG1uMetAlaValG1uThrValYlsG1n 263

RESULT 14
US-10-156-761-10012
; Sequence 10012, Application us/10156761
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHISA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10012
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10012

Alignment Scores:
Pred. No.: 4,31e-15 Length: 312
Score: 291.00 Matches: 84
Percent Similarity: 50.53% Conservative: 58
Best Local Similarity: 29.89% Mismatches: 99
Query Match: 8.14% Indels: 40
Db: 6 Gaps: 9

US-09-513-151-3 (1-2041) x US-10-156-761-10012 (1-312)
QY 83 CTTGAGAGGATTCGCGGGGACGCGGACCGCAATTCACGAGCTGCGGCTTGACGCTAGGC 142
Db 10 Val1IeAlaValAlG1YrPrOThrAla1a1aG1y1SerAspLeuG1yValPheLeuAla 29
QY 143 CAGCGGCTCGCGGGAGATGATGCTGAGCGCTACATGCATGCAGGTCGTATGAAGGCTTAGAC 202
Db 30 G1nArG1uLeuAspG1yG1uValAla1a1aSn1a1aSpMetG1nLeuYrArG1yMetAsp 49
QY 203 ATCATCACCAACAGAGTTCTGCCCAAGACAGAGAAATCTGCCGCGCACCAATGATGACG 262
Db 50 1IleG1Yr1ThrAla1YlLeuThrPrOg1uG1u1nArG1yG1y1IlePrOHis1Sh1eLeuAsp 69
QY 263 TTTGGATCTCTCTTGACCAATTCACAGACTGGTGACTTCAGAAATAGAGCAACTGCT 322
Db 70 1IleTrpAspVal1ThrValAla1a1a1aSerVal1a1aG1uYrG1nArG1uLeuAla1aArG1a 88

```

```

QY 323 CTGATGAGATATATATTCGCCGAGACAAATTCCTATTGTTGGGAGCAACATATAT 382
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 89 ArgIILeaSPAlaLeuLeuAlaIaGluGlyArgTTPProIleuValaIcYIleSerGlyLeu 108
QY 383 TACATGTAATCTCTGCTGTGGAAGATTCTTTCATATCCAGACCCGAGAGATGCGCAT 442
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 109 TyValArgGlyAlaValaIaSPanLeu-----GluPheProGlyThr 122
QY 443 GAGAAAGTATGACCGAAAGTGCAGCTTGAAAGAGGATGCTT----- 490
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 123 ASPProGluVal-----ArgAlaArgLeuGluGluGluLeuAlaLeuArgGlyProGly 140
QY 491 GATCTTCACAAAGCCTAGAGAGTGCAGCAAGAAATGCGTCCAGATGCATCCACT 550
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 141 AlAlaLeuHISAlaArgLeuAlaAlaAlaSPProGluAlaGlyHISAlaLeuSerSer 160
QY 551 GACAAAGCAGAAAGTGCAGAGCTTGCAAGTTTGAAGAAAGAGAAATCTCTACTAGT 610
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 161 AsnGlyArgArgIleValaArgAlaLeuGluValaIleGluIleThrGly----- 176
QY 611 GAATTTCTCATGCTCAACATACGAGAAAGTGTGCTCCCTTGAGAGTCCCTGGAAG 670
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 177 -----ArgProPheThrAlaAsn-----LeuProGlyHISAspSer 188
QY 671 TTCTTAACCCCTTGATCCTTGGCTTCATGCTGACGAGCAAGTTAGATGAGCGCTTG 730
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 189 ValTYrAspThrValaGlnIleGlyValaSPValaAlaArgProGluLeuAspGluArgGly 208
QY 731 GATTAAGAGGTGATGATGATGCTTGTGCTGAGGCTTGGAGAACTAGAGATTTTCAC 790
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 209 AlaArgArgValaSPArgMetTrpGluAlaGlyLeuValaSPLeuValaArg----- 225
QY 791 AGACGCTATATACAGAAAGATGTTTGGAAATAGCCAGACTATCAACATGTT----- 844
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 226 -----AlaLeuGluAlaGlnGlyLeuArgGluGluArgThr 237
QY 845 ANCTTCATCAATGATGCTTCAAGAAATTCACGAGTACTGATCATCAGGAGAAATGC 904
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 238 AlaSerArgAlaLeuGlyArgGlnGlnValaLeuAlaLeu-----AlaGlyGluCys 255
QY 905 ACA 907
    |||
DB 256 Thr 256
    |||

RESULT 15
PCT-US02-32727-18388
; Sequence 18388, Application PC/TUS0232727
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer
; APPLICANT: Skelky, Yasir
; APPLICANT: Persing, David
; APPLICANT: Bhatia, Ajay
; APPLICANT: Maisonneuve, Jean Francois
; APPLICANT: Zhang, Yann
; APPLICANT: Wang, Siqing
; APPLICANT: Jen, Shyian
; APPLICANT: Lodes, Michael
; APPLICANT: Benson, Darin
; APPLICANT: Jones, Robert
; APPLICANT: Carter, Darrick
; APPLICANT: Douglass, John
; TITLE OF INVENTION: Compositions and Methods for the Therapy and Diagnosis of Acnes V
; FILE REFERENCE: 210121, 514C1
; CURRENT APPLICATION NUMBER: PCT/US02/32727
; NUMBER OF SEQ ID NOS: 30992
; SEQ ID NO 18388
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Propionl acnes
PCT-US02-32727-18388

```

Alignment Scores:

Pred. No.:	7,05e-15	Length:	357
Score:	288.50	Matches:	81
Percent Similarity:	50.89%	Conservative:	62
Best Local Similarity:	28.83%	Mismatches:	95
Query Match:	8.07%	Indels:	43
DB:	1	Gaps:	8

```

US-09-513-151-3 (1-2041) x PCT-US02-32727-18388 (1-357)
QY 77 CTACCTCTTGATGATATCTCGGGCCAGCGGACCGCAAAATTCACGCTGCGTGGAG 136
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 47 LeuProValIleCysLeuValcIlyProThrAlaSerGlyIleSerGlyLeuAlaValArg 66
QY 137 CTAGCCACGCGCTC-----GGCGGTGATGCTGACGCGCTACTCATG 181
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 67 ValCysArgArgLeuTyValaSPGlnIleSProAlaGluIleAlaSPnThrAspMet 86
QY 182 CAGGCTATGAGAGGCTAGCATATCATACCAAGGTTTCCGCAAGCAGAGAAATC 241
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 87 ValValTYrArgGlyMetAspIleGlyThrAlaThrProThrLeuArgGluGlnArgThr 106
QY 242 TCCCGCACACATGATCATGCTTGTGATCCTCTTGACCAATTCACAGTGTGAGC 301
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 107 ValValHISHisLeuValSerIleLeuAspValThrValProSer-----SerLeuValLeu 125
QY 302 TTCAGAAATAGCAATCTGCTGATTAAGATATATTTGCCGAGACAAATTCCTAT 361
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 126 MetGlnThrLeuAlaArgSPAlaValaGluAspCysLeuSerArgGlyValIleProVal 145
QY 362 GTTGGGAGGAGCAATATATATCATGATCATCTGCTGCAAGGCTTGTGATATAC 421
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 146 LeuValaIcGlySerAlaLeuTyThrIlyAlaIleIleIaSPGluMetSerIleProPro 165
QY 422 AAGCCCGAGAGATGGCACTGAGAAAGTATGACCGAAAGTGCAGCTTGAAGAGAG 481
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 166 ThrAspProGluValaArgAlaArg-----TrpGlnGluIleuSPAlaGlu 181
QY 482 GATGCTCTTGATCTTACAAAGCCCTAAGCCAGGTGAGCCAGAAATGCTGCCAAGCTG 541
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 182 GlyProArgValaLeuHISAspGluLeuAlaArgAspProIlySPAlaIaGluSerIle 201
QY 542 CATCCATGACAAAGCAGAAAGTGCAGAGCTTGCAAGTTTGAAGAAAGAGAAATC 601
    |||... |||... |||... |||... |||... |||... |||... |||... |||...
DB 202 LeuProGlyAsnGlyArgArgIleValaArgAlaLeuGluValaIleSPLeuThrGlySer 221
QY 602 TCTCATATGTAATTT-----CTCATCTGTCACATACGAGAGAGGTGTGT 649
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 222 PheThrAlaThrIleProAspGlyThrLeuHISTrpProIlySThrValaGlnMetGly 240
QY 650 CCCCTTGAGAGTCCCTGTAAGTTCTCTAACCCTTGATCCTTGGCTTCATGCTGACAG 709
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 241 -----LeuIleuSerArg 245
QY 710 GCAGTCTAGATAGAGCGCTTGATAGAGGTGATGATGATGCTGCTGAGCTGCTG 769
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 246 LysAspIleAspGlnArgIleAlaSPArgValaSPGlnMetTrpAlaTyArgIlyPheVal 265
QY 770 GAGGAACATAGAGATTTTCACAGAGCCTAATACAGAGAAATGTTTGGAAAAATGACCG 829
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 266 AspGluValaArgSerLeu-----AlaAsnValaGly----- 275
QY 830 GACTATCAACATGATATCTTCATCA-----ATTGCTTCAAGAAATTTTCACAGATAC 883
    ... |||... |||... |||... |||... |||... |||... |||... |||...
DB 276 ---LeuArgGluGlyLeuThrAlaSerArgAlaLeuGlyTYrArgGlnValaLeuGluTy 294
QY 884 CTG 886
    |||
DB 295 Leu 295
    |||

```

Search completed: April 21, 2003, 18:58:58  
Job time : 99.4244 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:51:16 ; Search time 312.711 Seconds

(Without alignments)  
8416.087 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575

Sequence: 1 CTGCCATPAGATGCGTCGCC.....TTTACACAGAAAAA 2041

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Dgapop 6.0	Dgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 9138288

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-O=/cgn2.1/USPTO.spool/US09513151/runat.15042003.141145.26426/app.query.fasta.1.2446  
-DB=pending.patents\_AA\_Main -QFMT=fastan -SUFFIX=n2p.rapm -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blissum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09513151.ecgn.1.1.422.0runat.15042003.141145.26426  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=10 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending\_Patents\_AA\_Main:\*

1:	/cgn2.6/ptodata/1/paa/PCURS.COMB.pep:*
2:	/cgn2.6/ptodata/1/paa/US06.COMB.pep:*
3:	/cgn2.6/ptodata/1/paa/US07.COMB.pep:*
4:	/cgn2.6/ptodata/1/paa/US08.COMB.pep:*
5:	/cgn2.6/ptodata/1/paa/US081.COMB.pep:*
6:	/cgn2.6/ptodata/1/paa/US082.COMB.pep:*
7:	/cgn2.6/ptodata/1/paa/US083.COMB.pep:*
8:	/cgn2.6/ptodata/1/paa/US084.COMB.pep:*
9:	/cgn2.6/ptodata/1/paa/US085.COMB.pep:*
10:	/cgn2.6/ptodata/1/paa/US086.COMB.pep:*
11:	/cgn2.6/ptodata/1/paa/US087.COMB.pep:*
12:	/cgn2.6/ptodata/1/paa/US088.COMB.pep:*
13:	/cgn2.6/ptodata/1/paa/US089.COMB.pep:*
14:	/cgn2.6/ptodata/1/paa/US090.COMB.pep:*
15:	/cgn2.6/ptodata/1/paa/US091.COMB.pep:*
16:	/cgn2.6/ptodata/1/paa/US092.COMB.pep:*
17:	/cgn2.6/ptodata/1/paa/US093.COMB.pep:*
18:	/cgn2.6/ptodata/1/paa/US094.COMB.pep:*
19:	/cgn2.6/ptodata/1/paa/US095.COMB.pep:*
20:	/cgn2.6/ptodata/1/paa/US096.COMB.pep:*
21:	/cgn2.6/ptodata/1/paa/US097.COMB.pep:*
22:	/cgn2.6/ptodata/1/paa/US098.COMB.pep:*
23:	/cgn2.6/ptodata/1/paa/US099.COMB.pep:*
24:	/cgn2.6/ptodata/1/paa/US100.COMB.pep:*
25:	/cgn2.6/ptodata/1/paa/US101.COMB.pep:*
26:	/cgn2.6/ptodata/1/paa/US102.COMB.pep:*
27:	/cgn2.6/ptodata/1/paa/US060.COMB.pep:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2253	63.0	467	1	PCR-US02-07826-153	Sequence 153, App
2	2253	63.0	467	24	US-10-097-340-153	Sequence 153, App
3	2231	62.4	465	1	PCR-US02-07826-151	Sequence 151, App
4	2231	62.4	465	24	US-10-097-340-151	Sequence 151, App
5	1074	30.0	422	1	PCR-US01-01239-1394	Sequence 1394, App
6	1074	30.0	422	1	PCR-US01-01349-586	Sequence 586, App
7	1074	30.0	222	1	PCR-US01-01349-784	Sequence 784, App
8	1074	30.0	222	21	US-09-764-853-586	Sequence 586, App
9	1074	30.0	222	21	US-09-764-853-784	Sequence 784, App
10	1074	30.0	222	21	US-09-764-902-1394	Sequence 1394, App
11	1074	30.0	222	24	US-10-072-326-586	Sequence 586, App
12	1074	30.0	222	24	US-10-072-326-784	Sequence 784, App
13	686.5	19.2	143	21	US-09-757-028-2199	Sequence 2199, App
14	686.5	19.2	143	26	US-10-222-911-2199	Sequence 2199, App
15	597.5	16.7	252	16	US-09-270-767-44570	Sequence 44570, App
16	549	15.4	430	19	US-09-513-151-2	Sequence 2, App11
17	548	15.3	428	21	US-09-791-537-6482	Sequence 46482, App
18	536.5	15.0	427	21	US-09-791-537-76494	Sequence 76494, App
19	471	13.2	461	21	US-09-708-427-19098	Sequence 19098, App
20	456	12.8	451	21	US-09-708-427-19099	Sequence 19099, App
21	403.5	11.3	314	21	US-09-791-537-70445	Sequence 70445, App
22	391	10.9	423	21	US-09-708-427-19100	Sequence 19100, App
23	389	10.9	306	15	US-09-107-532A-5131	Sequence 5131, App
24	389	10.9	306	15	US-09-107-532A-5131	Sequence 17815, App
25	386.5	10.8	386	16	US-09-248-796-17815	Sequence 17815, App
26	386.5	10.8	386	27	US-60-096-409-17815	Sequence 17815, App
27	367	10.3	257	25	US-10-155-881-8448	Sequence 8448, App
28	355	9.9	329	21	US-09-710-279-1460	Sequence 1460, App
29	339	9.5	319	12	US-08-821-326-3288	Sequence 3288, App
30	339	9.5	319	20	US-09-611-529-4328	Sequence 4328, App
31	338	9.5	312	21	US-09-708-427-27588	Sequence 27588, App
32	338	9.5	318	21	US-09-708-427-27587	Sequence 27587, App
33	334.5	9.4	311	17	US-09-368-382-32	Sequence 32, App1
34	334.5	9.4	311	20	US-09-649-145-32	Sequence 32, App1
35	334.5	9.4	311	26	US-10-260-877-32	Sequence 32, App1
36	333.5	9.3	329	19	US-09-513-996A-62990	Sequence 62990, App
37	327.5	9.1	290	19	US-09-513-996A-62991	Sequence 62991, App
38	324	9.1	308	21	US-09-708-427-27589	Sequence 27589, App
39	320	9.0	576775	12	US-08-895-611-2	Sequence 2, App11
40	320	9.0	576775	12	US-08-895-611D-2	Sequence 2, App11
41	320	9.0	576775	18	US-09-458-180-2	Sequence 2, App11
42	320	9.0	576775	22	US-09-895-611D-2	Sequence 2, App11
43	319.5	8.9	329	19	US-09-513-996A-18307	Sequence 18307, App
44	319.5	8.9	342	15	US-09-198-452A-983	Sequence 983, App
45	319.5	8.9	345	18	US-09-438-185-912	Sequence 912, App

#### ALIGNMENTS

RESULT 1

Sequence 153, Application PC/US0207826

GENERAL INFORMATION:

APPLICANT: Millennium Pharmaceuticals, Inc. et al.

TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, Assessment, Prevention, and Therapy of Ovarian Cancer

FILE REFERENCE: MRI-030PC

CURRENT APPLICATION NUMBER: PCT/US02/07826

PRIOR FILING DATE: 2002-03-14

PRIOR APPLICATION NUMBER: 60/276,025

PRIOR FILING DATE: 2001-03-14

PRIOR APPLICATION NUMBER: 60/325,149

PRIOR FILING DATE: 2001-09-27

PRIOR APPLICATION NUMBER: 60/276,026

```

PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 153
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-07826-153

Alignment Scores:
Pred. No.      1,03e-210      Length:      467
Score:         2253.00      Matches:      439
Percent Similarity: 94.22%      Conservative: 1
Best Local Similarity: 94.00%      Mismatches: 1
Query Match:    63.02%      Indels:      26
DB:             1          Gaps:         1

US-09-513-151-3 (1-2041) x PCT-US02-07826-153 (1-467)
QY 11 ATGGCGTCCGTCGCGGTCGACGACGATTCCTGTGGCAGTGGGCTCAGGGGCTGCAA 70
DB 1 Metlaaervlaalaalaalargalavalprovalglgyserygleuhalgylgleucln 20
QY 71 CGAGCCCTACCTCTGTAGTATTCCTGCGGGCCACGGGCGACGGCAATCCACGCTGCG 130
DB 21 Argthrleuproleuvalvalilleuuglyalathrglylthrglylserthrleuala 40
QY 131 TTGACGTAGCGCCGCGGCTGGGGTGAGATCTCAGCGCTGACTCATCAGGTCTAT 190
DB 41 Leuclnleuglylnargyleuylgylglnlevalserlaaasrmetglnvaltyr 60
QY 191 GAAGCGCTAGACATCATCACCAACAGGTTCTGCCCAAGACGACAGATCTGCCGCGAC 250
DB 61 Gluclnleuaspriellelthaslnlyvalserlaaglnuglnahgyllecysarghls 80
QY 251 CACATGATCAGCTTGTGGATCTCTTGTGACAAATTAACAGTGGTGGACTTCAGAAAT 310
DB 81 Hismetliserphevalasprroleuvalthrasnlytrthvalvalasprheargasn 100
QY 311 AGACCACTGCTGTGATGGAAGATATATTTGCCGAGCAAAATTCCTATTGTGGGA 370
DB 101 Argalaathralaleuillleglnasprillephealaargaspysilleproillevalglly 120
QY 371 GAACCAATTTATACATTAATCTGCTGTGAAAAGTTCTGTCAATACCAAGCCCGAC 430
DB 121 Glythrlnsnlytrylleglnserleuclnleuvalasnthlylserprocln 140
QY 431 GAGATGGGACGAGAAAGTGTGACCGAAAGTGGAGCTTGAAGAGAGATGCTTT 490
DB 141 Gluclnleuglylthrglylvalilleasprargylvalglnleuclnlysglnasprglyeu 160
QY 491 CTACTTCACAAAGCCCTTAAGCCAGGTGACCCCAAAATGGTGGCAAGCTCATCCACAT 550
DB 161 Valclnleuhslyargyleuclnvalasprroglumetalalalyseuclnlsprchls 180
QY 551 GACAAGCAAGTGGCGAGAGCTTCAGATTTTGAAGAAACAGGATCTCATAGT 610
DB 181 Aspysargylvalalaargserleuclnvalaphegluglnthrglylleserhisser 200
QY 611 GAATTTCCATCGTCAACATACGAAAGAGTGGTGGTCCCTTGAGAGTCCCTGTAAG 670
DB 201 Gluclnleuhslyargyleuclnlsprthrglylgllylserleuclnlysglnleucln 220
QY 671 TTCTCAACCTGCTGCTTGGCTTCATGCTGACAGGAGCTTCAGATGAGACGCTG 730

```

```

DB 221 PheSerlnsprocysilleleutrlpneuhsalaasprlnlaalvalleuasprgluargleu 240
QY 731 GATPAAGAGGGTGATGACATGCTGTGCTGGGCTGTGAGAGCAATPAAGATTTTTCAC 790
DB 241 Aspysargylvalaasprmetleuclnlaalagllyleuclnlysglnleuclnlysglnleucln 260
QY 791 AGAGCGTATATATCGAAGATGTTTCGGAATATGCGGACGATTCACATGATGATCTG 850
DB 261 Argargtyrasnclnlysasvalserclnlnasnsersclnsptryrglnhslylephe 280
QY 851 CAATCAATTTGGCTTCAGAGAAATTCACGAGTACATGATGATGAGGAGAAATGACACTG 910
DB 281 Gluclnleuglylthrglylserleuclnlysglnleuclnlysglnleuclnlysglnleucln 300
QY 911 GAGACTAGTAAACGCTTCTAAGAAAGA----- 940
DB 301 Gluclnleuhslyargyleuclnlysglylleglnlaaleuclnlysglnvalthrglyarg 320
QY 941 -----CCTGGTCCCAT 952
DB 321 TyrAlaarglyslasnaargtrpvallylsasnaargpheleuSerArgproglyprolle 340
QY 953 GTCCCCCTGTCTATGCTTGAAGTATCTGATGCTGCAAGTGGAGAGTGTCTT 1012
DB 341 ValProvalaalytrglyleuclnlysglnleuclnlysglnleuclnlysglnleucln 360
QY 1013 GAACCTGCTCTTGAATCGTCAAAAGTTTCATCCAGGCGCCAAAGCTTACAGCCATCCA 1072
DB 361 Gluclnleuclnlysglnleuclnlysglnleuclnlysglnleuclnlysglnleucln 380
QY 1073 ATAAAGTCCATATCAATGAGTGTGACACAGAGAAATATACCTGTGTGACCTGTGT 1132
DB 381 Ileuclnleuclnlysglnleuclnlysglnleuclnlysglnleuclnlysglnleucln 400
QY 1133 GATGCAATCATCATTTGGGGATGCGCAATGGGCGACGCGCAATCAATCCACATCT 1192
DB 401 Asparglylletllelleglyasprarglultrpalaalansllylserlyserhsleu 420
QY 1193 AACCACTGAGAAAGAAAGAAAGTGGAGTCAATGCTGCTCAACACCATGAAGTCAAG 1252
DB 421 Asnclnleuclnlysglnleuclnlysglnleuclnlysglnleuclnlysglnleucln 440
QY 1253 AGTGTTCCTCCAGACTATACAAAGAACTTAAGGAGGATGCCAGGCGACAGATGAT 1312
DB 441 Servalserproasprlnaslnlysglnleuclnlysglnleuclnlysglnleucln 460
QY 1313 CAAGAGCTGAAGTCAAGCTT 1333
DB 461 Gluclnleuclnlyscyserval 467

RESULT 2
US-10-097-340-153
Sequence 153, Application US/10097340
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOERSCHE
APPLICANT: Shubhang KAMATKAR
APPLICANT: Steve G KOVATS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Aml SEN
APPLICANT: Peter VERBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, JR.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumel ZHOU
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MKI-030

```

```

CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 153
LENGTH: 467
TYPE: PRT
ORGANISM: Homo sapiens
US-10-097-340-153

Alignment Scores:
Pred. No.: 1,03e-210 Length: 467
Score: 2253.00 Matches: 439
Percent Similarity: 94.22% Conservative: 1
Best Local Similarity: 94.00% Mismatches: 1
Query Match: 63.02% Indels: 26
DB: Gaps: 1

US-09-513-151-3 (1-2041) x US-10-097-340-153 (1-467)

OY 11 AGGCGTCGCTGGGCTGCACAGAGATTCCTGTGGAGTGGGCTGAGGCGCTGCAA 70
DB 1 MetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArgGlyLeuGln 20
OY 71 CGAGCCCTGCTCTTGTAGTATTCYGGGGCCAGCGGACCGGCAATTCACGCTGGCG 130
DB 21 ArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSerThrLeuAla 40
OY 131 TTGCAGTAAAGCCAGCGGCTGGGCTGAGATGTCAGCCCTGACTCCATGCAAGTCTAT 190
DB 41 LeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValSerIleAspSerMetGlnValTyr 60
OY 191 GAAGGCTGACATCATCAACCAAGAGTTTCGCCCAAGAGAGCAAGAACTCCGCGCAC 250
DB 61 GlnGlyLeuAspIleIleThrAsnLysValSerIleGlnIleGlnArgIleCysArgHis 80
OY 251 CACATATGACGTTTGTGATCCTCTGTGACCAATTCACAGTGTGACTTCAGAAAT 310
DB 81 HisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAspPheArgAsn 100
OY 311 AGAGCACTGCTGTGAATGAATATATTTGCCCGAGACAAATTCCTATTGTGTGGGA 370
DB 101 ArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIleValIleGly 120
OY 371 GGAACCAATATATACATTTGAATCTGCTGCTGAAAGTTCTTGTCATACCAAGCCCGAG 430
DB 121 GATTTAAsnTyrTyrIleGlnSerLeuLeuTrpLysValIleuValAsnThrLysProGln 140
OY 431 GAGATGGGCACTGAGAAAGATGTGACCAAGAGTGGAGCTGAAAGAGAGATGCTTT 490
DB 141 GlnMetGlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGlnAspGlyLeu 160
OY 491 GTAATTCACAACGCTTAAGCAGGTGACCCAGAAATGGCTGCCAAGTGCATCCAT 550
DB 161 ValLeuHisLysArgLeuSerGlnValAspProGlnMetAlaIleLysLeuHisProHis 180
OY 551 GAAACGCAAAAGTGGCCAGAGCTTCCAGATTTTGAAGAAACAGAAATCTCCATAGT 610
DB 181 AspLysArgLysValAlaArgSerLeuGlnValPheGlnGlnThrGlyIleSerHisSer 200

OY 611 GAATTTCTCCATGCTCAACATACGGAAGAAGTGGTGGTCCCTTGAGGCTCTGAGG 670
DB 201 GlnPheLeuHisArgGlnHisThrGlnGlnGlyGlyProLeuGlyLeuProLeuLys 220
OY 671 TTCTCTAACCCCTTGCTGGCTTGCATGGCGACGAGGAGCTTGTATGAGGCTGG 730
DB 221 PheSerAspProCysIleLeuTrpLeuHisAlaAspGlnAlaValLeuAspGlnArgLeu 240
OY 731 GATTAAGAGGTGATGACATGCTTGGCTGGCTTGTGAGGAACTAAGATTTTCAC 790
DB 241 AspLysArgValAspSerPheLeuAlaIleGlyLeuLeuGlnGlnLeuArgAspPheHis 260
OY 791 AGAGCTATATATCAGAAGATGTGGGAAATAGCCAGACATATCAATGATGATCTTC 850
DB 261 ArgArgTyrAsnGlnLysAsnValSerGlnAsnSerGlnAspTyrGlnHisGlyIlePhe 280
OY 851 CAACTCAATGGCTTCAAGGAATTCACGAGTACCTGATCAGTGAAGGAAATGACACTG 910
DB 281 GlnSerIleGlyPheLysGlnPheHisGlyTyrLeuIleThrGlnGlyLysCysThrLeu 300
OY 911 GAGACTAGTAACCGCTTCAAGAAAGAA-----CTGGTCCCAT 940
DB 301 GlnThrSerAsnGlnLeuLeuLysGlyIleGlnAlaLeuLysGlnValThrLysArg 320
OY 941 -----CTGGTCCCAT 952
DB 321 TyrAlaArgLysGlnAsnArgTyrValLysAsnArgPheLeuSerArgProGlyProIle 340
OY 953 GTCCCCCTGCTATGCTGATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1012
DB 341 ValProValTyrGlyLeuGlnValSerAspValSerLysTrpGlnGlnSerValLeu 360
OY 1013 GAACCTGCTCTGAAATCTGCAAGTTTCATCCAGGCGCAAGGCTTACAGGCTACCA 1072
DB 361 GlnProAlaLeuGlnIleValGlnSerPheIleGlnGlyHisLysProThrAlaThrPro 380
OY 1073 ATAAGATGCCATATCAATGAGCTGAGACAAAGAAATATACCTGTGTGACCTCTGT 1132
DB 381 IleLysMetProTyrAsnGlnGlnIleGlnValAspLysArgSerTyrHisLysCysAspLeuCys 400
OY 1133 GATCGAATCATCAATGGGAGTTCGGAATGGGCGGACGACATTAATCCAAATCCCACTTG 1192
DB 401 AspArgIleIleIleGlyAspArgGlnTrpAlaAlaHisIleLysSerLysSerHisLeu 420
OY 1193 AACCACTGAAGAAAGAAAGAAATTTGACTCAGATGCTGTCAACACCATAGAAGTCA 1252
DB 421 AsnGlnLeuLysLysArgArgArgLeuAspSerAspAlaValAsnThrIleGlnSerGln 440
OY 1253 AGTGTTCCTCCAGACTATTAACAAAGAACTTAAGGGAAGGATCCCGAGGCAAGATAT 1312
DB 441 SerValSerProAspHisAsnLysGlnProLysGlnLysGlySerProGlnGlnAsnAsp 460
OY 1313 CAAGAGCTGAATGCAAGCTGT 1333
DB 461 GlnGlnLeuLysCysSerVal 467

RESULT 3
PCT-US02-07826-151
Sequence 151, Application PC/TUS0207826
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030PC
CURRENT APPLICATION NUMBER: PCT/US02/07826
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14

```



```

: PRIOR APPLICATION NUMBER: 60/324,967
: PRIOR FILING DATE: 2001/09/26
: PRIOR APPLICATION NUMBER: 60/311,732
: PRIOR FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: 60/325,102
: PRIOR FILING DATE: 2001-09-26
: PRIOR APPLICATION NUMBER: 60/323,580
: PRIOR FILING DATE: 2001-09-19
: NUMBER OF SEQ ID NOS: 363
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 151
: LENGTH: 465
: TYPE: PRT
: ORGANISM: Homo sapiens
PCT-0502-07826-151

Alignment Scores:
Pred. No.: 1,47e-208      Length: 465
Score: 2231.00           Matches: 437
Percent Similarity: 93.79%   Conservative: 1
Best Local Similarity: 93.58% Mismatches: 1
Query Match: 62.41%         Indels: 28
                        Gaps: 2

US-09-513-151-3 (1-2041) x PCT-US02-07826-151 (1-465)

QY 11 ATGGCGTCGGTGGCGGCTGCACGAGCATTCCTGTGGGAGTGAGGCTCAGGGCCCTGCA 70
Db 1 MetAlaserValAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 20

QY 71 CGGACCTACTCTTTGATGATATTTCTGGGGCCAGGGGACCGGCAATCCAGCTGGCG 130
Db 21 ArgThrLeuProLeuValAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 38

QY 131 TTGAGCTAGGCGCGCGCTGGCGGGGAGAGATGCTGACGCGTGCATCCATGACAGCTAT 190
Db 39 LeuAlnLeuGlyAlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 58

QY 191 GAAGCGCTAGACATCATCACCAACAGGTTTCTGCCAAGACGAGAAATCTGCCGCGAC 250
Db 59 GluGlyLeuAspIleIleThrAlnAlnValSerAlaAlnGluAlnAlAlAlAlAlAlAlAl 78

QY 251 CACATGATCAGCTTTGTGATCCTCTTGTGACCAATTACACAGTGTGACTTCAGAAAT 310
Db 79 HisMetIleSerPheValAspProLeuValAlThrAsnTyrAlnAlAlAlAlAlAlAlAl 98

QY 311 AGACAACCTGCTGATTTGAAGATATATTTGGCCGAGACAAATTCCTATATTGTGGGA 370
Db 99 ArgAlaThrAlaLeuAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 118

QY 371 GGAACCAATTTATACATTTGAATCTCTGCTGTGAAAGTTCTTGCAATACCAAGCCGAC 430
Db 119 GlyThrAsnTyrTyrIleGluSerLeuLeuThrPylAlAlAlAlAlAlAlAlAlAlAlAl 138

QY 431 GAGATGGGCACTGAGAAAGTGATTTGACCGGAAAGTGAGGCTTGAAAAGAGATGCTTT 490
Db 139 GluMetGlyThrGluValAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 158

QY 491 GTACTTCACAAAGCCCTAAGCGAGAGTGGACCCGAAATGGCTGCGCAAGCTGCATCCACAT 550
Db 159 ValLeuAlnIleAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 178

QY 551 GACAACGCAAAAGTGGCCAGAGCTTGCAAGTTTGTGAAAGAAACAGAAATCTGTACATG 610
Db 179 AspArgAlaGlyValAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 198

QY 611 GAATTTTCTCATGCTGCAACATPAGGAAAGAAAGTGAGTGGTCCCTTGAGGCTCTGCAAG 670
Db 199 GluPheLeuAlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 218

QY 671 TTCTCTAACCCTTCACATCTTTGGCTTCAGATGCTGACAGGCACTTTCAGATGAGCGTTG 730
Db 219 PheSerAlnProCysIleLeuThrPheAlnAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAlAl 258

```

QY 731 GATPAGAGGGGTGGATGACATCGCTGCTGCTGGGCTTTGGAGCAACTAGAGATTTTCC 790  
 Db 239 ASPLYSATGVALASPRASPRMCTLEUALAALAGLYLENDLCLNGLULEULATGASPRPHENIS 258  
 QY 791 AGACGGCTATAATCAGAAAGATGTTTTCGAAATATGCGCAGAGACTATACATGCTATCTTC 850  
 Db 259 ATGATGTYTASNGLNDLYSASNDVALSERGLUASNSERGLASPRYRGLNHISGLYLERPHE 278  
 QY 851 CAATCAATTGGCTTCAAGGAATTTTCACGAGTACCTGATCATCTGAGGAAATGACACATCG 910  
 Db 279 GINSERTLEGLYPHELYSGILUPHENISGLUTRYLEULEITLTHCLNGLYLSCYSFTRLEU 298  
 QY 911 GAGACTAGTAACACCGCTTTCTAAAGAGA----- 940  
 Db 299 GUTTHRSERASNGLNDLEULYSLYSGLYILEGLUALALEULSGLINALTHRLYSARG 318  
 QY 941 -----CTGTGCCATT 952  
 Db 319 TYRLLATARGLYSGLASNAATGTRVALLYLSASNAARGPHELEUSERARGPROGLPROLLE 338  
 QY 955 GTCCCCCGCTGTATAGGCTTTAGAGGTATGTGATGTCGCAAGTGGGAGAGATCTGTCTT 1012  
 Db 339 VALPROFOVALTYGLYGLYENGLUVALSERASVALSERLYSTRPLNGLUSERVALLEU 358  
 QY 1013 GAACCTGCTTGGAAATCGTGCAGAAATTTTCATCCSAGGGCCACAGCCTACAGCATCCCA 1072  
 Db 359 GUPTROALALENGLILLEVALINSERPHLEIENGILYHISLYSPROTHRLATPRPO 378  
 QY 1073 ATAAGATGCCATACATGATAGCTGAGACAGAGAGAGATTCACCTGTGTGACCTGTG 1132  
 Db 379 IELYSMERPTOTYASNDLUALAGLUAANLYSARGSETYHISLEUCYSASPRLEUCYS 398  
 QY 1133 GATCGAATCATATTGGGGATGCGGANTGGCAGCGCACATAAATCCAATCCCACTTG 1192  
 Db 399 ASPARGILEILEILEGLYSPARGIUTRPAIALAHISILEYSESLYSSERHISLEU 418  
 QY 1193 AACCAACTGAAAGAAAGAGATTGAGCTGAGATGCTGCAACCCATAGAAATCG 1252  
 Db 419 ASNGLNDLEULYLSYATGTATGATGLEUASPRSERPALAVALASMTHTLLEGLUSERGLN 438  
 QY 1253 AGTGTTCGCCAGATATATACAAAGACCTTAAGGAGAGGATCCCCAGGCGAGAATGAT 1312  
 Db 439 SERVALSERPROASPHISASNDLYSGIUNPROLYSGILYGLYSETRPROLYGLNASHASP 458  
 QY 1313 CAAGAGCTGAATGCAAGCCTT 1333  
 Db 459 GINGLULEULYSCYSESERVAL 465  
 RESULT 4  
 US-10-097-340-151  
 ; Sequence 151, Application US/10097340  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John MONAHAN  
 ; APPLICANT: Manjula GANNANARAPU  
 ; APPLICANT: Sebastian HOERSCH  
 ; APPLICANT: Shubhangi KAMATKAR  
 ; APPLICANT: Steve G. KOVATS  
 ; APPLICANT: Rachel E. MEYERS  
 ; APPLICANT: Michael MORRISSEY  
 ; APPLICANT: Peter OLANDT  
 ; APPLICANT: Ami SEN  
 ; APPLICANT: Peter VEIBY  
 ; APPLICANT: Gordon B. MILLS  
 ; APPLICANT: Robert C. BAST, Jr.  
 ; APPLICANT: Karen LU  
 ; APPLICANT: Rosemarie SCHMANDT  
 ; APPLICANT: Xumei ZHAO  
 ; APPLICANT: Karen GLATT  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification  
 ; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
 ; FILE REFERENCE: MRI-030  
 ; CURRENT APPLICATION NUMBER: US/10/097, 340



```

; COMMENT FILING DATE: 2001-VI-14
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939

```

APPLICANT: human genome sciences, inc., et al.

```

: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PUZ06PCT
: CURRENT APPLICATION NUMBER: PCT/US01/01349
: CURRENT FILING DATE: 2001-01-14
: Prior application data removed - consult PAM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: Patentl Ver. 2.0
: SEQ ID NO 784
: LENGTH: 222
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (124)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (145)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: PCT-US01-01349-784

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 1 Gaps: 0

US-09-513-151-3 (1-2041) x PCT-US01-01349-784 (1-222)
OY 2 TGCCATAAGATGCGTCGCGGCGCTGCACGAGCATTCCTGTGGCAGTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCTGCGACGAGACCCCTACCTTGTAGTGAFTCTCGGGGCCACGGGCAAAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACGCTGGCGTTGCACCTAGCGCGCGGCTGGAGATCGTCAGCGCTCAGTCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlyIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAAGGCTAGACATCATCAACAAGGTTTCCGCCAAGAGCAGAAATC 241
Db 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
OY 242 TGGCGGACACCATATGACGTTTGTGATCTCTTGTGACCAATTAACAAGTGGTGC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAlaAsp 100
OY 302 TTCAGAAATAGCAACAGCTGCTGATGAGATATATTGCCCGAGCAAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGTGGAGAGAACCAATATTATACATGAATCTGCTCTGGAAGTTCTTGTCAATAC 421
Db 121 ValValIgly***ThrAsnTyrTyrIleGlnSerLeuLeuTrrPlyValLeuValAsnThr 140
OY 422 AAGCCCCAGAGATGGGACATGAGAAAGTGAACCCGAAAAGTGGAGCTTGAAGAAGAG 481
Db 141 LysProGlnGln***GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGln 160
OY 482 GATGCTCTTGTACTCACAACCGCTAGCGCGAGTGGAGCCGAGAAATGGCTGCCAAGTC 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetCAlaIalysLeu 180
OY 542 CATCCACATGACAAACGAAAGTGGCCAGAGCTTGCAGATTGTTGAAGAAACAGAAATC 601
Db 181 HisProHisAspLysArgLysValAlaIalArgSerLeuGlnValPheGlnGlnIleThrGlyIle 200
OY 602 TCTCATAGTGAATTTCTCCATCGTCACATAC-GGAAGAAGGTGCTGCTCCCT 654
Db 201 SerHisSerGlnPheLeuHisArgGlnHisThrGlyArgLysTrrPrrSerPro 218

```

```

RESULT 8
US-09-764-853-586
: Sequence 586, Application US/09764853
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PUZ06
: CURRENT APPLICATION NUMBER: US/09/764,853
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - consult PAM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: Patentl Ver. 2.0
: SEQ ID NO 586
: LENGTH: 222
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (124)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (145)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: US-09-764-853-586

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 21 Gaps: 0

US-09-513-151-3 (1-2041) x US-09-764-853-586 (1-222)
OY 2 TGCCATAAGATGCGTCGCGGCGCTGCACGAGCATTCCTGTGGCAGTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
OY 62 GGCTGCGACGAGACCCCTACCTTGTAGTGAFTCTCGGGGCCACGGGCAAAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValAlaIleLeuGlyAlaThrGlyThrGlyLysSer 40
OY 122 ACGCTGGCGTTGCACCTAGCGCGCGGCTGGAGATCGTCAGCGCTCAGTCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlyIleValSerAlaAspSerMet 60
OY 182 CAGGTCTATGAAGGCTAGACATCATCAACAAGGTTTCCGCCAAGAGCAGAAATC 241
Db 61 GlnValTyrGlnGlyLeuAspIleIleThrAsnLysValSerAlaGlnGlnArgIle 80
OY 242 TGGCGGACACCATATGACGTTTGTGATCTCTTGTGACCAATTAACAAGTGGTGC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAlaAsp 100
OY 302 TTCAGAAATAGCAACAGCTGCTGATGAGATATATTGCCCGAGCAAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
OY 362 GTTGTGGAGAGAACCAATATTATACATGAATCTGCTCTGGAAGTTCTTGTCAATAC 421
Db 121 ValValIgly***ThrAsnTyrTyrIleGlnSerLeuLeuTrrPlyValLeuValAsnThr 140
OY 422 AAGCCCCAGAGATGGGACATGAGAAAGTGAACCCGAAAAGTGGAGCTTGAAGAAGAG 481
Db 141 LysProGlnGln***GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGln 160
OY 482 GATGCTCTTGTACTCACAACCGCTAGCGCGAGTGGAGCCGAGAAATGGCTGCCAAGTC 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetCAlaIalysLeu 180
OY 542 CATCCACATGACAAACGAAAGTGGCCAGAGCTTGCAGATTGTTGAAGAAACAGAAATC 601

```

```

Db 181 HisProHisAspLysArgValAlaArgSerLeuGlnValPheGluIuThrGlyIle 200
QY 602 TGTCTAGTAGAATTTCTCCATGCTCAACATAC-GGAGAAGGTGGTCCCT 654
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTrpIlePro 218

RESULT 9
US-09-764-853-784
; Sequence 784, Application US/09764853
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1206
; CURRENT APPLICATION NUMBER: US/09/764,853
; PRIOR APPLICATION DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 784
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-784

```

```

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 21 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-764-853-784 (1-222)

```

QY 2 TGCATTAAGATGCGCTCCGCGGCTGCACGAGCAGTTCCTGTGGCAGTGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGCCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGCCACGGGCAACATCC 121
Db 21 GylLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 ACGCTGGCGCTTGCAGTACGAGCGGCTCGGCGGTGAGATGCTCAGCCTCAGTCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValSerIleAspSerMet 60
QY 182 CAGGTCTATGAAGGCTTACATCATCATCACCAAGGTTCTGCCCAAGAGAGAAATCC 241
Db 61 GlnValIlyrGlyGlyLeuAspIleIleThrAsnLysValSerAlaGlnIuInArgIle 80
QY 242 TGCCGGCACCATATGATCAGCTTGTGATCCTCTTGACCAATTCACAGTGTGAC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnIlyrThrValValAsp 100
QY 302 TTCAGAAATAGCAACTGCTGTGATGAAGATATATTGCCCGAGACAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGAGAGAACCAATTATACATTCGATCTGCTGTGACCAATTCCTCAATACC 421
Db 121 ValValGly***ThrAsnIlyrIleGlnSerLeuLeuTrpLysValLeuValAsnThr 140
QY 422 AAGCCCCAGAGATGGGCACTGAGAAAGTATGACCGAAAGTGGAGCTTGAAGAGAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGlnLeuGlnLysGln 160

```

```

QY 482 GATGGCTTGTACTTTCACAAAGCCTAAGCGAGTGGACCCAGAAATGGCTGCCAAGTG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGlnMetValAlaLysLeu 180
QY 542 CATCCACATGACAAACGCAAGTGGCCAGAGCTTGCAGATTGTTGAACAACAGCAATC 601
Db 181 HisProHisAspLysArgValAlaArgSerLeuGlnValPheGluIuThrGlyIle 200
QY 602 TGTCTAGTAGAATTTCTCCATGCTCAACATAC-GGAGAAGGTGGTCCCT 654
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTrpIlePro 218

```

# RESULT 10

```

US-09-764-902-1394
; Sequence 1394, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1213
; CURRENT APPLICATION NUMBER: US/09/764,902
; PRIOR APPLICATION DATE: 2001-01-17
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1394
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-902-1394

```

```

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 21 Gaps: 0

```

US-09-513-151-3 (1-2041) x US-09-764-902-1394 (1-222)

```

QY 2 TGCATTAAGATGCGCTCCGCGGCTGCACGAGCAGTTCCTGTGGCAGTGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValGlySerGlyLeuArg 20
QY 62 GGCCTGCAAGGAGCCCTACCTCTGTAGTATCTCGGGCCACGGGCAACATCC 121
Db 21 GylLeuGlnArgThrLeuProLeuValValIleLeuGlyAlaThrGlyThrGlyLysSer 40
QY 122 ACGCTGGCGCTTGCAGTACGAGCGGCTCGGCGGTGAGATGCTCAGCCTCAGTCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlyGlnArgLeuGlyGlyGlnIleValSerIleAspSerMet 60
QY 182 CAGGTCTATGAAGGCTTACATCATCATCACCAAGGTTCTGCCCAAGAGAGAAATCC 241
Db 61 GlnValIlyrGlyGlyLeuAspIleIleThrAsnLysValSerAlaGlnIuInArgIle 80
QY 242 TGCCGGCACCATATGATCAGCTTGTGATCCTCTTGACCAATTCACAGTGTGAC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnIlyrThrValValAsp 100
QY 302 TTCAGAAATAGCAACTGCTGTGATGAAGATATATTGCCCGAGACAAATTCCTATT 361
Db 101 PheArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGAGAGAACCAATTATACATTCGATCTGCTGTGACCAATTCCTCAATACC 421

```

```

Db 121 ValValGly***ThrasnTyrTrlIleGluSerLeuLeuTrpLysValLeuValAsnThr 140
QY 422 AAGCCCGAGAGATGGGCACTGAGAAAGTATGACCGGAAAGTGGAGCTTGAAGAGAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGluLeuGlnLysGlu 160
QY 482 GATGGCTCTTGACTTCACAAAGCGCTTAAGCCAGGTGGAGCCCGAAAGTGGCTCCAAAGCTG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeu 180
QY 542 CATCCACATGACAAAGCGAAAGTGGCCAGGAGCTTGCAGAGTTTGAAGAAACAGGAATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGlnThrGlyIle 200
QY 602 TCTCATAGTGAATTTCTCCATCGTCACATAC-GGAAGAAGGTGGTGTCCCT 654
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTrpTrpSerPro 218

RESULT 11
US-10-072-326-586
; Sequence 586, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206C1
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See File wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 586
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-326-586

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 24 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-072-326-586 (1-222)
QY 2 TGGCATAAATGAGCGTCCGCGGCTGCACGACGAGTTCCTGGGCACTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValIlySerGlyLeuArg 20
QY 62 GGCCCTCAACGAGACCTTCTTGTAGTATCTCGGGCCGCGGACCGGCAACGCAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValIleLeuGlnYalAlaThrGlyThrGlyLysSer 40
QY 122 AGCCTGGCGTGGAGCTAGGCGGCGGCTCGGAGTGAATCGTCGACGCTGACTCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlnArgLysGlnArgLysGlyGluIleValSerAlaAspSerMet 60
QY 182 CAGGTATATGAAGCGCTTAACATCATCATCACCACAAAGTTTCTGCCCAAGGACGAATC 241
Db 61 GlnValTyrGlnGlyLeuAspIleThrAsnLysValSerAlaGlnGlnGlnArgIle 80
QY 242 TGGCGGACACCATGATCAGCTTGTGATCCCTGTGTGACAAATTAACAGAGGTGAGAC 301
Db 81 CysArgHisHisMetIleSerPheValAspProLeuValThrAsnTyrThrValValAsp 100

```

```

QY 302 TTCAGAAATAGACAACTGCTCTGATGTAGATATATTTGCCCGAGCAAAATTCCTAT 361
Db 101 PheAlaArgAsnArgAlaThrAlaLeuIleGlnAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGGGAGAGAACCAATATTACATTTGATATGCTGTCTGTGAAAGTTCTTGTCAATAC 421
Db 121 ValValGly***ThrasnTyrTrlIleGluSerLeuLeuTrpLysValLeuValAsnThr 140
QY 422 AAGCCCGAGAGATGGGCACTGAGAAAGTATGACCGGAAAGTGGAGCTTGAAGAGAG 481
Db 141 LysProGlnGlu***GlyThrGlnLysValIleAspArgLysValGluLeuGlnLysGlu 160
QY 482 GATGGCTCTTGACTTCACAAAGCGCTTAAGCCAGGTGGAGCCCGAAAGTGGCTCCAAAGCTG 541
Db 161 AspGlyLeuValLeuHisLysArgLeuSerGlnValAspProGluMetAlaIleLysLeu 180
QY 542 CATCCACATGACAAAGCGAAAGTGGCCAGGAGCTTGCAGAGTTTGAAGAAACAGGAATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGlnThrGlyIle 200
QY 602 TCTCATAGTGAATTTCTCCATCGTCACATAC-GGAAGAAGGTGGTGTCCCT 654
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTrpTrpSerPro 218

RESULT 12
US-10-072-326-784
; Sequence 784, Application US/10072326
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P206C1
; CURRENT APPLICATION NUMBER: US/10/072,326
; CURRENT FILING DATE: 2002-02-11.
; Prior Application removed - See File wrapper or Palm
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 784
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-072-326-784

Alignment Scores:
Pred. No.: 2,06e-95 Length: 222
Score: 1074.00 Matches: 214
Percent Similarity: 98.62% Conservative: 1
Best Local Similarity: 98.17% Mismatches: 3
Query Match: 30.04% Indels: 1
DB: 24 Gaps: 0

US-09-513-151-3 (1-2041) x US-10-072-326-784 (1-222)
QY 2 TGGCATAAATGAGCGTCCGCGGCTGCACGACGAGTTCCTGGGCACTGGGCTCAGG 61
Db 1 CysHisLysMetAlaSerValAlaAlaAlaArgAlaValProValIlySerGlyLeuArg 20
QY 62 GGCCCTCAACGAGACCTTCTTGTAGTATCTCGGGCCGCGGACCGGCAACGCAATCC 121
Db 21 GlyLeuGlnArgThrLeuProLeuValIleLeuGlnYalAlaThrGlyThrGlyLysSer 40
QY 122 AGCCTGGCGTGGAGCTAGGCGGCGGCTCGGAGTGAATCGTCGACGCTGACTCCATG 181
Db 41 ThrLeuAlaLeuGlnLeuGlnArgLysGlnArgLysGlyGluIleValSerAlaAspSerMet 60
QY 182 CAGGTATATGAAGCGCTTAACATCATCATCACCACAAAGTTTCTGCCCAAGGACGAATC 241

```

```

Db 61 GlnValTylGluLysLeuAspIleIleThrAsnLysValSerIleGlnGluAsnIle 80
QY 242 TGGCGGACACCATGATCAGCTTGTGATCCTCTGTGACCAATACAGTGTGGAC 301
Db 81 CysAlaHISHisMetIleSerPheValAspProLeuValThrAsnTyrThrValAsp 100
QY 302 TTCAGAAATAGACACTGCTGATGATGAAGATATATTGGCCGACAAATAATCTTAT 361
Db 101 PheAlaAsnHrValAlaThrAlaLeuIleGluAspIlePheAlaArgAspLysIleProIle 120
QY 362 GTTGTGGAGGAGCAATATTACATTGAATCTCTGCTGAAAATCTTGTCAATACC 421
Db 121 ValValIly**ThrAsnTyrTyrIleGluSerLeuLeuTyrPylValLeuValAsnThr 140
QY 422 AAGCCCGAGAGATGGGACCTGGAAGAGTATTGACCCGAAAGTGGAGCTTGAAGAG 481
Db 141 LysProGlnGlu**GlyThrGluLysValIleAspArgLysValGluLeuGluLysGlu 160
QY 482 GATGCTCTGTACTTACAAAGCCTAGCAGAGTGGACCCAGAAATGGCTGCACAGCTG 541
Db 161 AspGlyLeuValLeuHisLysAlaGluSerGlnValAspProIleMetAlaLysLeu 180
QY 542 CATCACATGACAAAGCAAGAGTGGCAGAGCTTGCAGATTTTTGAAGAAACAGAGATC 601
Db 181 HisProHisAspLysArgLysValAlaArgSerLeuGlnValPheGluGluThrGlyIle 200
QY 602 TTCATAGAGGAATTTTCCATCTGCTCAACATAC-GGAAGAGGTGGTGGCTCCCT 654
Db 201 SerHisSerGluPheLeuHisArgGlnHisThrGlyArgLysTyrPheSerPro 218

```

## RESULT 13

```

US-09-513-151-3 (1-2041) x US-09-757-028-2199 (1-143)
; Sequence 2199, Application US/09757028
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM001
; CURRENT APPLICATION NUMBER: US/09/757,028
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 2660
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2199
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-028-2199

```

## Alignment Scores:

```

Pred. No.: 1,47e-57 Length: 143
Score: 686.50 Matches: 131
Percent Similarity: 97.10% Conservative: 3
Best Local Similarity: 94.93% Mismatches: 3
Query Match: 19.20% Indels: 1
DB: 21 Gaps: 1

```

US-09-513-151-3 (1-2041) x US-09-757-028-2199 (1-143)

```

QY 920 AACCAAGCTTCTAAAGAAAGAGACCTGTCCATTGCCCCCTGTATGCTTAGAGTA 979
Db 7 AsnArgPheLeuSerArg---ProGlyProIleValProProValTyrGlyLeuGluVal 25
QY 980 TCTGATGCTCGAAGTGGAGAGTCTGTTTGAACCTGCTTGAATCGTGAAGT 1039
Db 26 SerAspValSerLysTyrIleGluLysValLeuGluProAlaLeuGluIleValGlnSer 45
QY 1040 TTCATTCAGAGGCGCACAGGCTACAGCCACTCCAAATAAAGATGCCATACATGAAGTGA 1099
Db 46 PheIleGlnGlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnLysGlu 65

```

```

QY 1100 AACCAAGAGATTATACCTGTGTGACTCTGTGATCGAATCATCATTTGGGATCGCGAA 1159
Db 66 AsnLysArgSerTyrHisLysLeuCysAspLeuCysAspArgIleIleIleGlyAspArgGlu 85
QY 1160 TGGCGGACGACATTAATAATCCAAATCCACTTGAACCACTGAAGAAAGAGAGATTG 1219
Db 86 TrpAlaAlaHisIleLysSerLysSerHisLysAsnGlnLeuLysLysArgArgGlu 105
QY 1220 GACTCAGATGCTGTCAACACATAGAAAGTGAAGTGTTCCTCCAGACTATACAAAGAA 1279
Db 106 AspSerAspAlaValAsnThrIleGluSerGlnSerValSerProAspHisAsnLysGlu 125
QY 1280 CCTAAAGGAGAGGATCCCGAGGCGCAATGATCAAGACTGAATGACAGCTT 1333
Db 126 ProLysGluLysGlySerProGlyGlnAsnAspGlnGluLeuLysCysSerVal 143

```

## RESULT 14

```

US-10-222-911-2199
; Sequence 2199, Application US/10222911
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM001CIN
; CURRENT APPLICATION NUMBER: US/10/222,911
; CURRENT FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: 09/757,028
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 2660
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2199
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-911-2199

```

```

Alignment Scores:
Pred. No.: 1,47e-57 Length: 143
Score: 686.50 Matches: 131
Percent Similarity: 97.10% Conservative: 3
Best Local Similarity: 94.93% Mismatches: 3
Query Match: 19.20% Indels: 1
DB: 26 Gaps: 1

```

US-09-513-151-3 (1-2041) x US-10-222-911-2199 (1-143)

```

QY 920 AACCAAGCTTCTAAAGAAAGAGACCTGTCCATTGCCCCCTGTATGCTTAGAGTA 979
Db 7 AsnArgPheLeuSerArg---ProGlyProIleValProProValTyrGlyLeuGluVal 25
QY 980 TCTGATGCTCGAAGTGGAGAGTCTGTTTGAACCTGCTTGAATCGTGAAGT 1039
Db 26 SerAspValSerLysTyrIleGluLysValLeuGluProAlaLeuGluIleValGlnSer 45
QY 1040 TTCATTCAGAGGCGCACAGGCTACAGCCACTCCAAATAAAGATGCCATACATGAAGTGA 1099
Db 46 PheIleGlnGlyHisLysProThrAlaThrProIleLysMetProTyrAsnGlnLysGlu 65
QY 1100 AACCAAGAGATTATACCTGTGTGACTCTGTGATCGAATCATCATTTGGGATCGCGAA 1159
Db 66 AsnLysArgSerTyrHisLysLeuCysAspLeuCysAspArgIleIleIleGlyAspArgGlu 85
QY 1160 TGGCGGACGACATTAATAATCCAAATCCACTTGAACCACTGAAGAAAGAGAGATTG 1219
Db 86 TrpAlaAlaHisIleLysSerLysSerHisLysAsnGlnLeuLysLysArgArgGlu 105
QY 1220 GACTCAGATGCTGTCAACACATAGAAAGTGAAGTGTTCCTCCAGACTATACAAAGAA 1279
Db 106 AspSerAspAlaValAsnThrIleGluSerGlnSerValSerProAspHisAsnLysGlu 125

```

OY 1280 CCTAAGGGAAGGATCCCGAGGCGAGATGATCAAGCTGAAATGACGCTT 1333  
 Db 126 ProlysglyglyserProglylnasnasprnglnleuLycyserval 143

# RESULT 15 US-09-270-767-44570

; Sequence 44570, Application US/09270767  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Homburger et al.  
 ; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 ; FILE REFERENCE: File Reference: 7326-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; NUMBER OF SEQ ID NOS: 62517  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 44570  
 ; LENGTH: 252  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila melanogaster  
 US-09-270-767-44570

## Alignment Scores:

Pred. No.:	1,01e-48	Length:	252
Score:	597.50	Matches:	119
Percent Similarity:	67.48%	Conservative:	47
Best Local Similarity:	48.37%	Mismatches:	77
Query Match:	16.71%	Indels:	3
DB:	16	Gaps:	2

US-09-513-151-3 (1-2041) x US-09-270-767-44570 (1-252)

OY 71 CGGACCTACCTCTGTGATGATCTCGGGCCAGCGGCGCAATTCACGCTGCG 130  
 Db 5 ArglyValProleuileValValleuGlyserThrGlyThrGlyThrLysleuser 24  
 OY 131 TTGCAGCTAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 190  
 Db 25 LeuGlnleuAlaGlnArgheGlyGlyGlnleuIleSerAlaAspserMetGlnValTyr 44  
 OY 191 GAAGGCTAGACATCATCACAACAGCTTCTGCCAGACAGAGCAATTCGCGGCAC 250  
 Db 45 ThrHisleuAspIleAlaThrAlaLysAlaThrLysGlnGlnSerArgAlaArgHis 64  
 OY 251 CACATGATCAGCTTTGTGGATCCTCTGTGACCAATTACACAGTGGGACTTCAGAAAT 310  
 Db 65 HisleuLeuAspValThrThrPro--AlaGlnProPheThrValThrHisPheArgAsn 83  
 OY 311 AGAGCACTGCTGTGATGATATATTTGCCCGAGCAAAATTCCTATGTGTGGCA 370  
 Db 84 AlaAlaLeuProIleValGlnArgleuAlaLysAspThrSerProIleValValGly 103  
 OY 371 GGAACCAATTATTAATGATGATCTGCTGGAAGTCTTGTGCAATAC-----AAG 424  
 Db 104 GlyThrAsnTyrTyrIleGlnSerleuLeuTrpAspIleleuValAspserAspValLys 123  
 OY 425 CCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAAGTGAAGCTTGAAGAGAGAT 484  
 Db 124 ProGlnGlnGlyLysProserGlyLysHisleuLysAspAlaGlnleuAsnAlaLeuSer 143  
 OY 485 GGTCTTGTTACTCACAACGCTTAAGCCAGGTGGACCCAGAAATGCGTCCAAAGCTGCAT 544  
 Db 144 ThrleuGlnleuHisGlnleuLeuAlaLysIleAspAlaGlySerAlaAsnArgIleHis 163  
 OY 545 CCACATGACAACGCAAAAGTGGCCAGAGCTTGCAAGTTTGAAGAACAAGCAATCTCT 604  
 Db 164 ProAsnAsnArgArgLysIleIleArgAlaIleGlnValTyrGlnSerThrGlyGlnThr 183  
 OY 605 CATAGTGAATTTCTCCATGCTCAACATAGGAAGAAGTGTGTGCTCCCTTGAGAGTCT 664  
 Db 184 LeuSerGlnMetleuAlaGlnArgAlaGlnProGlyLysAsnArgleuGlyGlyPro 203  
 OY 665 CTGAAGTTCTTAACCTTGACCTCTTGGCTTCATGCTGACCGAGGAGTTCTAGATGAG 724

Db 204 LeuArgTyrProHisIleValleuLeuTrpLeuArgCysGlnGlnAspValleuAsnGln 223  
 OY 725 CCGTTGATTAAGAGGAGGTGGATGACATGCTTGTGCTGGGCTCTTGGAGAACTAGAGAT 784  
 Db 224 ArgleuAspSerArgValAspGlyMetleuAlaGlnGlnleuLeuProGlnleuArgGln 243  
 OY 785 TTTCACAGACGCTAAT 802  
 Db 244 PheHisAsnAlaHisHis 249

Search completed: April 21, 2003, 19:12:17  
 Job time : 321.711 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OW nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 18:41:35 ; search time 22.0286 seconds

(without alignments)  
5452.195 Million cell updates/sec

Title: US-09-513-151-3

Perfect score: 3575  
Sequence: 1 CTGCCTAAGATGCGCTCCG.....TTTACAGAAAAA 2041

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Delop 6.0, Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=rat  
-Q=/cgn2\_1/USFTO\_spool/US09513151/runat\_15042003\_141144\_26406/app\_query.fasta.1.2446  
-DB=Issued\_Patents\_AA -OFMT=fasta -SUFFIX=n2p.rat -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09513151 -GCN 1.1 23 -eunat\_15042003\_141144\_26406 -NCPU=6 -ICPU=3  
-NO\_XLPRY -NO\_MMAP -LARGEOUTRY -NBS\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents, AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6C.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/6D.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307.5	8.6	193	US-09-134-001C-4171	Sequence 4171, Ap
2	114.5	3.2	563	US-09-134-001C-3172	Sequence 3172, Ap
3	99	2.8	1024	US-09-562-737-50	Sequence 50, Appl
4	99	2.8	1548	US-08-463-092B-7	Sequence 7, Appl
5	99	2.8	1548	US-08-460-907B-7	Sequence 7, Appl
6	98.5	2.8	1044	US-08-777-405A-2	Sequence 2, Appl
7	98.5	2.8	1044	US-08-977-871A-2	Sequence 2, Appl
8	98.5	2.8	1044	US-09-225-951A-2	Sequence 2, Appl
9	97	2.7	511	US-09-019-201A-2	Sequence 2, Appl
10	96.5	2.7	430	US-09-310-363C-6	Sequence 6, Appl
11	95.5	2.7	3052	US-08-557-122A-26	Sequence 26, Appl
12	95.5	2.7	3052	US-09-262-666-26	Sequence 26, Appl

13	94	2.6	529	US-08-836-442-3	Sequence 3, Appl
14	94	2.6	566	US-09-040-725A-1	Sequence 1, Appl
15	94	2.6	739	US-08-836-943-2	Sequence 2, Appl
16	94	2.6	859	US-09-199-637A-281	Sequence 281, Appl
17	93.5	2.6	236	US-09-605-858-34	Sequence 5, Appl
18	93.5	2.6	240	US-09-073-587-5	Sequence 5, Appl
19	93	2.6	674	US-08-961-083-200	Sequence 200, Appl
20	93	2.6	633	US-08-235-836C-72	Sequence 72, Appl
21	92.5	2.6	544	US-08-559-397A-29	Sequence 29, Appl
22	92.5	2.6	3075	US-08-460-309-5	Sequence 5, Appl
23	92.5	2.6	3075	US-08-125-077-5	Sequence 5, Appl
24	92.5	2.6	3696	US-09-134-001C-5080	Sequence 5080, Appl
25	91.5	2.6	976	US-09-104-324B-4	Sequence 4, Appl
26	91.5	2.6	2482	US-08-328-254-6	Sequence 6, Appl
27	91	2.5	1150	US-09-002-285-74	Sequence 74, Appl
28	90.5	2.5	429	US-09-074-512-1	Sequence 1, Appl
29	90.5	2.5	510	US-09-134-001C-3368	Sequence 3368, Appl
30	90.5	2.5	649	US-09-134-001C-3891	Sequence 3891, Appl
31	90.5	2.5	693	US-08-235-836C-68	Sequence 68, Appl
32	90.5	2.5	1201	US-09-098-901-2	Sequence 2, Appl
33	90.5	2.5	1872	US-08-188-582-14	Sequence 14, Appl
34	90.5	2.5	1872	US-08-646-715-14	Sequence 14, Appl
35	90.5	2.5	1893	US-08-188-582-11	Sequence 11, Appl
36	90.5	2.5	1893	US-08-646-715-11	Sequence 11, Appl
37	90	2.5	425	US-09-310-363C-8	Sequence 8, Appl
38	90	2.5	474	US-09-134-001C-4953	Sequence 4953, Appl
39	90	2.5	630	US-09-360-545-78	Sequence 78, Appl
40	90	2.5	655	US-08-264-002-2	Sequence 2, Appl
41	90	2.5	1384	US-08-976-255-11	Sequence 11, Appl
42	89.5	2.5	1262	US-09-357-251-33	Sequence 33, Appl
43	89.5	2.5	1266	US-08-468-557-4	Sequence 4, Appl
44	89.5	2.5	1266	US-09-357-251-32	Sequence 32, Appl
45	89.5	2.4	2511	US-09-261-907-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-134-001C-4171  
Sequence 4171, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NOCLEDIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: GTC-007  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4171  
LENGTH: 193  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4171

#### Alignment Scores:

Pred. No.: 1.52e-25  
Score: 307.50  
Percent Similarity: 61.80%  
Best Local Similarity: 36.52%  
Query Match: 8.60%  
DB: 4  
Gaps: 3

US-09-513-151-3 (1-2041) x US-09-134-001C-4171 (1-193)

QY 68 CAAGGACCCCTACTCT--CTTGATGATTCGCGGCGACGCGCAATCCAGC 124  
DB 3 GlnMetThrLysPhePheLeuIleValIleValIleGlyProThrAlaSerIleLysThrIle 22



```

Db 365 -----Ileasplileglyllysleusenglnluarqargayalaalarggluleu 382
OY 1163 GCAGCCGACATAAATCCAAATCCACTTGACCAACTGAAGAAAAG----- 1210
Db 383 Argasphsilevalsergluileglnasleuglnmetlysasphlaasleugluile 402
OY 1211 -----AGAAGATTGACCTCAGATGCTGTCACACACATAGAAAGTCAGAGT-----GTT 1258
Db 403 Serphelysproleusargluprothrllegllyllegluphevalglupheleuile 422
OY 1259 TCCCCAGACTATACAAAGAACTAAAGGAGGATCCCA 1300
Db 423 Serproasn-----Argglglupro 429

RESULT 3
US-09-562-737-50
; Sequence 50, Application US/09562737
; Patent No. 6428967
; GENERAL INFORMATION:
; APPLICANT: Herz, Joachim
; APPLICANT: Gotthardt, Michael
; TITLE OF INVENTION: LDL Receptor Signaling Pathways
; FILE REFERENCE: UTS0708
; CURRENT APPLICATION NUMBER: US/09/562,737
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 1024
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-562-737-50

Alignment Scores:
Pred. No.: 0.0913 Length: 1024
Score: 99.00 Matches: 84
Percent Similarity: 40.16% Conservative: 63
Best Local Similarity: 22.95% Mismatches: 159
Query Match: 2.77% Indels: 60
DB: 4 Gaps: 19

US-09-513-151-3 (1-2041) x US-09-562-737-50 (1-1024)
OY 293 GTGTGAGCTTACAAATATAGCAACTGCTGATGAGATATATTATTCGCGAGACAA 352
Db 101 Metvalaspraserlysthrvalthrasmleumethrlllecysalaarglilegln 120
OY 353 ATTCTATTGTTGTGGAGAGAACCAATTATTAATGATCTGCTCTGGAAGTTCTT 412
Db 121 Ile-----Thrasmhlsasprolytserleuargatgslulemet 134
OY 413 GTCAATATCCAAAGCCCGAGATGGGCACT-----GAGAAAGTGATT----- 454
Db 135 glulglulysylsgerluserlilethrglythrleuargllyasphlyvalleuileuargasp 154
OY 455 GACCGAAAGTAGGAGCTGAAAGAGGATGCTTGTACTCTGC-----AAGCGCTA 508
Db 155 :::::::::::::::::::::leuhssthraspraptyrleu 171
OY 509 AGCCAGGTGACCCAGAAATGGCTGCCAAG-----CTGCATCCACATGACAAACGC 559
Db 172 Asntrpleuasprhlsigllyatgthrilaarglulglulnlyalglulglulnhsiglulasp 191
OY 560 AAAGTGCCACAGAGCTGCAAGTTTGTGAAGAAACAGAGAACTCTCATAGTGAATT--- 616
Db 192 leuileuargllyasphethyrgluasphlnasvalaspraserargaspptrophegln 211
OY 617 CTCACATGTCACATACAGGAAGAGGTGGTGGCTTGGAGGT-----CCTCTGAG 670
Db 212 leuasnleuileuargllyalglulnlagllyaspraplleuasnnglyserhlsprthlsier 231

```

```

OY 671 TTCTTACCCCTTGACCTCTTGGCTTCATGCTGACACGACGACTTCTAGATGACGCTTG 730
Db 232 Pheasplysalacysgluphealaallephelglnysglulnleghlprolylsasn 251
OY 731 GATAGAGGGGTGATGACATGCTGCTGCGGGCTCTGGAGAGACATGAGATTTTTCAC 790
Db 252 glulglulys-----Hlsylsalgllyphenet---aspheulysasphleu 266
OY 791 AGACGCTATATCAGAGAAATGTTTCGGAATAACCCAGGACATCAACAT----- 841
Db 267 Prolysglnasvallysglnlysglulargllylleglnlnlalnhslysaasnycys 286
OY 842 GGTATCTTCCAAATATGCTTCCACAGAAATTTACAGAGTACATGATCATGAGGAAA 901
Db 287 Glylnmetarglulleglnalalysvalargllyalserleuallaargserleuyls 306
OY 902 TGCACACTGAGACTAGTACACAGCTCTCAAGAA-----GGACCTGGTCCCAT 952
Db 307 ---Thryrgllythrserpheleuvallysglnlysmetvalgllylsasnlyslu 325
OY 953 GTCCCGCTCTATGCTTATGAGGTATCTGATGCTCGAATGCTCGAGAGAGCTGTTCTT 1012
Db 326 Valproargleuargllylilethrllysglucysvalmetargllyasphlylly 343
OY 1013 GAACCTGCTTGAATCGTCAAGATTTC-----ATCCAGGCCACAAAGCTTACAGCC 1066
Db 344 -----Thrysglulvalileglnlalatrpasnlleuthrasnlelylsargltyraspala 361
OY 1067 ACTCCATTAAGATGCCATCAATGAGCTGAGACAGAAAGAGTTTACCTGCTGAC 1126
Db 362 Serprolyserphenrleuasphlyllyasptryglinspelyltyrtyrserphegln 381
OY 1127 CTCTGTGAT-----CGAATCATCTTGGGATCGGATCGGATGGAGCCGAC 1171
Db 382 Thrthrleuglulglulnlealaglyleuilealaglyllyrleasr---llellehs 400
OY 1172 ATAAATCCAAATCCCACTTGACCAACTGAAGAAAGAAAGATTTGACTGAGTGTCT 1231
Db 401 Lyslysllyslysser-----Lysasphlsphelileuaglulglulysp 415
OY 1232 GTCACACCATAGAAAGTACAGAGTGTTCCTCA-----GAC 1267
Db 416 Gluserthrmetlysgluasprervalserprolylsylserleuvalleuglnngln 435
OY 1268 TATAACAAGAACTTAA 1285
Db 436 Tytrasmargllymetlys 441

RESULT 4
US-08-463-092B-7
; Sequence 7, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B

```

```

FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steed, Carol Miernicki
REGISTRATION NUMBER: 39,539
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
US-08-463-092B-7

Alignment Scores:
Pred. No.: 0.12 Length: 1548
Score: 99.00 Matches: 74
Percent Similarity: 36.63% Conservative: 52
Best Local Similarity: 21.51% Mismatches: 88
Query Match: 2.77% Indels: 130
DB: 1 Gaps: 19

US-09-513-151-3 (1-2041) x US-08-463-092B-7 (1-1548)
OY 14 GCGTCGCGGCGGCGTCGACAGCAGTCTGTCGCGGCGGCGG-----CTCAGGGC 64
DB 1271 AAGGlyAlaAlaSerSerAlaLeuHisProValGlnAlaGlySerLeuValLeuGly 1290
OY 65 CTGCAA-----CGGACCCCTACCTCTTGTAGTG----- 91
DB 1291 ValGlnMetArgTyrArgGlnGlyLeuProLeuValLeuArgGlyValSerPheGlnIle 1310
OY 92 -----ATTCTCGGGCCACGGGCGACCGGCAATCCACGCTGGCG 130
DB 1311 AlaProAlaGlnLysValGlyValGlyValGlyArgTyrGlnGlySerGlyThrLeuLeu 1330
OY 131 TTGACAGTAGGCGACGCGGCGCGGTGAGATCGTCAGCGGTGACTGCATGAGGTAT 190
DB 1331 LeuThrPhe-----MetArgMetVal 1337
OY 191 GAA-----GGCCTAGACATCATCAAC-----AAGTTTCTGCCCAAGACAGAGA 238
DB 1338 GluValLysGlyValIleHisValAsnGlyArgGlnMetSerAlaTyrGlyLeuArg 1357
OY 239 ATCTCGCGGCGACACATGATGCTTGTG-----GATCCTCTGTGACCAATTACACA 292
DB 1358 GluLeuArgTyrGlnHis---PheSerMetIleProGlnAspProValLeuPheAspGlyThr 1376
OY 293 GTG----- 295
DB 1377 ValAlaGlnAsnValAspProPheLeuGlnAlaSerSerAlaGlnValIleTrrAlaLeu 1396
OY 296 -----GTGGACTTCAGAAATAGACCAACTGCTGTGATTGAGATATATTGCCCGAGAG 349

```

```

DB 1397 GluLeuValGlyLeuArgGlnArgValAlaSerGluSerGlnGlyIleAspSerArg----- 1415
OY 350 AAAATTCATTATGTTGGGAGGAGCAACATTATATGTAATGTAATCTGCTGTGAAAGTT 409
DB 1416 -----ValLeuGlnGlyGlySerAsnTyrSerVal----- 1425
OY 410 CTTGCAATACCAAGCCCGAGAGATGGGCACTGAGAAAGTCATGTGACCGAAAGTGAG 469
DB 1426 -----GlyGlnArgGlnLeuMetCysMetAlaArgAlaLeu----- 1437
OY 470 CTTGAAAGAGAGATGCTCTGTACTT---CACAAACCCCTAAGCAGAGTGACCCAGAGA 526
DB 1438 LeuLysArgGlySerGlyPheIleLeuMetAspGlnAlaThrAlaAsnIleAspProAla 1457
OY 527 ATGCGTCCCAAGCTCATCCACATGACACAACGCAAAAGTGCCGAGAGCTTGCAAGTTT 586
DB 1458 Leu-----AspArgGlnIleGlnAlaThrValMetSerAlaPhe 1470
OY 587 GAGAAACAGAAATCTCTATGTAATTTCTCCATCGTCAACATACGGAAGAGTGT 646
DB 1471 -----SerAlaTyrThrValIleThrIleAlaHisArgLeuHisThrVal----- 1485
OY 647 GGTCCCTTGAGAGTCTCTGAGAGTTCTTAACCTTGCACTTGCCTTCATCTGAC 706
DB 1485 ----- 1485
OY 707 CAGGCACTTATGATGACGCGCTTGATAGAGCGGTGATGACATGCTGCTGGGCTC 766
DB 1486 -----AlaGlnTyrAspIleIle-----ValMetAspIleGlyVal 1498
OY 767 TTGAGAGACTAAGAGATTTTCACAGACGCTTATTCAGAAAGATGTTTGGAAATAGC 826
DB 1499 ValAlaGlnMetGly-----SerProArgGlnLeuValMetAsnHis 1512
OY 827 CAGGACTATCAACATGATCTTCCATCAAT-----GGCTCAAGAAATTTGAC 877
DB 1513 GlnSerMetPheHisSerMetValGlnSerLeuGlnSerArgGlySerLysAspPheTyr 1532
OY 878 GAGTACTGATC 889
DB 1533 GluLeuLeuMet 1536

RESULT 5
US-08-460-907B-7
Sequence 7, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
TITLE OF INVENTION: RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424

```







REFERENCE/DOCKET NUMBER: 27866/33441

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 474-6300

TELEFAX: (312) 474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1044 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-225-951-2

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
Score: 98.50	1044	96	66	174	137
Percent Similarity: 34.25%					
Best Local Similarity: 20.30%					
Query Match: 2.76%					

US-09-513-151-3 (1-2041) x US-09-225-951-2 (1-1044)

```

QY 291 CAGTGGTGACTTCAGAAATAGAGCAACTGCTGATTGAGATATATTGGCCGAGACA 350
DB 10 GTPHETPTHTLysgluInsnInserValValValAspPheLeuLeuProThrgly 29
QY 351 AATTCCTA-----TTGTTG 365
DB 30 ValTyrLeuAsnPheroValSerArgAsnAlaAsnLeuSerThrIleLysgluLeu 49
QY 366 TGGAG-----GACCAATTATTACATTGATCTGCTGCTGAGAACT-----408
DB 50 TTPHSAArgAlaGlnTyrGlnPro-LeuPheHisMetLeuSerGlyProGluAlaTyrVa 69
QY 409 -----TC 410
DB 69 LPhetHrcysIleAsnGlnThrAlaGluGlnGluLeuGluAspGluGlnArgyle 89
QY 411 TTGTCATACCA-----GCCCGAGAGATGGCA 440
DB 89 uCysAspValGlnProPheLeuProValLeuArgValAlaArgGluIAsp-Argy 109
QY 441 CTGAGAAAGTGTATGACCAAGAGAGTGTGAAAGAGAGAGTGTGTTGACTTACA 500
DB 109 aLlysLysLeuIleAsnSerGlnIleSerLeuLeuIleGlyLysGly-----LeuHisG 127
QY 501 AAGGCTAAGCCAGGTGAGCCAGAAATG-----GCTCCAGATGTCATCCACATG 551
DB 127 LurPheAspSerLeuCysAspProGluValAlaAsnAspPheArgAlaLysMetCysGlnPhec 147
QY 551 ACAAAAGCAAGTGGCCAGAGAGCTTGCAGTTTGAAGAAACAGAGATCTCATAGTG 611
DB 147 ySgluGluAlaAlaArgArgGlnGlnGlyTrpGluAlaTrpLeuGlnTyrSerP 167
QY 612 AATTTCCTCATGCTCAACTAGCAGAGAGAGTGGTGGCTTGGAGGCTCTGTAAT 671
DB 167 hPheProLeuGlnLeuGlnProSerAlaGlnThrTrpGlyPro-----GlyThrLeuArgL 185
QY 672 TCTTAACCTTGCATCTT-----TGCTTCATGCTGACGAGGAGTCTAG 719
DB 185 euProAsnArgAlaLeuLeuValAsnValLysPheGluLysSerGluGlnSerPheThP 205
QY 720 ATGAGCGCTTGATAGAGGCTGATGACATCTTGCTGCT-----760
DB 205 heGlnValSerThrLysAspValProLeuAlaLeuMetAlaCysAlaLeuArgLysLysA 225
QY 761 -----GGCTCTTGAGAGAACTAAGATTTTCAACAACGCTAATC 803
DB 225 lArThValPheArgGlnProLeuValGlnGlnProGluAspTyrThrLeuGlnValAsnG 245
QY 804 AGAAGATGTTTCGAGAAATATACCACTATCAACATGTATCTTCCATCATCAATGGCT 863

```

```

DB 245 LyrGHis-----GluTyrLeuTyrGlyAsnTyrProLeuCysGlnP 259
QY 864 TCAAGAA-----TTTACAGACTACTGTATCTAGGAGAAATATGACACTGG 911
DB 259 heGlnTyrIleCysSerCysLeuHisSerGlyLeuThrProHis-----LeuThrMetV 277
QY 912 AGACTAGTACCAGCTTCA-----AAGAAAGACCTGGTCCCATGTC- 955
DB 277 aLHisSerSerIleLeuAlaMetArgAspGluGlnSerAsnProAlaProGlnValG 297
QY 956 -----CCCCCTGTATAGCTTGAAGTATGATCTGATCTGCAAGT 995
DB 297 InLysProArgAlaLysProProPheProAlaLysLysProSerSerValSerLeuT 317
QY 996 GGGAGAGTCTGTTCTTGAACCTGCTCTGAAATCTGCAAACTTTCATCCAGGCCACA 1055
DB 317 rPser-----LeuGluGlnProPheArgyle-----GluLeuIleGlnGlySerL 332
QY 1056 AGCTACAGCCACTCCA-----ATAAGATGCCATACAGTGAAGCTGAGAACAGAAAGTT 1112
DB 332 ySValAsnAlaAspGluArgMetLysLeuValValGlnAlaGlyLeuPheHisGlyAsnG 352
QY 1113 ATCACTGT-----GACCTGTGTATGCAATCATCATTTG 1148
DB 352 lMetLeuCysLysThrValSerSerGlnValSerValCysSerGlnProVal-----370
QY 1149 GGGATGCGCAATGGGAGCGGCACATTAATCCAAATCCCACTTGAAACCACTGAAAGAAA 1208
DB 371 -----TrpLysGlnArgLeuGlnPheAspIleAsnIleCysAspLeuProArgM 387
QY 1209 GAAGAGATTGAGCTCAGATGCTGCACACACATAGAAAGTGTGTTCCCGAGACT 1268
DB 387 eAlaIArgLeuCysPheAlaLeuTyrAlaValIleGluLysAlaLysLysAlaArgSerT 407
QY 1269 ATAACAAGAACCTAAGAGGAGGATCCCGAGGAGATGATCAAGACTGGAATGCA 1328
DB 407 hLysLysLysSerLysLys-----413
QY 1329 GCCTTTAAGAGACATGTCAGTGGCTTGGAAAGTGTGGGAGATTCAGTTCAGG-----1384
DB 414 -----AlaAspCysProIleAlaIleAlaIleAlaIleAlaIleLeuPheAspTyrLysAspG 431
QY 1385 -----AGGAGGGGATGTTGTCCTCCAGTGG 1414
DB 431 InLeuLysThrGlyGluArgCysLeuTyrMetT 442

```

RESULT 9  
US-09-019-201A-2  
Sequence 2, Application us/09019201A  
Patent No. 5968780

GENERAL INFORMATION:  
APPLICANT: FENG, PING  
APPLICANT: SOPPER, DANIEL R.  
APPLICANT: LI, YI  
APPLICANT: DILLON, PATRICK J.  
TITLE OF INVENTION: DENDRITIC CELL-DERIVED GROWTH FACTOR  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: HUMAN GENOME SCIENCES, INC.  
STREET: 9410 KEY WEST AVENUE  
CITY: ROCKVILLE  
STATE: MD  
COUNTRY: US  
ZIP: 20850

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/019,201A  
FILING DATE: HEREWITH  
CLASSIFICATION: 536



```

ATTORNEY/AGENT INFORMATION:
NAME: A. ANDERS BROOKES
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PF346
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 511 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-019-201A-2

Alignment Scores:
Pred. No.: 0.0961 Length: 511
Score: 97.00 Matches: 75
Percent Similarity: 37.73% Conservative: 48
Best Local Similarity: 23.01% Mismatches: 106
Query Match: 2.71% Indels: 97
DB: Gaps: 16

US-09-513-151-3 (1-2041) x US-09-019-201A-2 (1-511)

QY 371 GGAAACAATATTACATGTAATCTGCTGGAAAGTTCTTCAATACCAAGCCAG 430
DB 25 GYSERALALEUSETILEASRGLUTRATGALANILEUDEULYUSGLYMETMET 44
QY 431 GAGATGGGACSTGAGAAATGATGACCCAAAGTGGCTTAAAGAGGATGGCTT 450
DB 45 ATGLEUGLYLARGLEUVALLEUASNTHLYSGULGULEUAAASNGUATGLEUMET 64
QY 491 GTACTTCACAAAGCCGTAAGCCAGTGGACCCAGAAATGGCTGCCAATGCCA 547
DB 65 THLEU-----LYSILEAGLUMETLYSGULAMETKRGTHLEUILLERPROPRO 82
QY 548 -----CATGACAAAGCCAAAGTGGCCAGAGCTTGCAAGTTT 586
DB 83 SERMETHISRPHERHEGLNALALYSNILEULLEGLUALRGER--GLNALPHEASNILE 101
QY 587 -----GAAACAAGGAAATCTGCTCATAGT 610
DB 102 LEUATGMEETPROLYSGILALALALEUHNISLEUHNISAPRILEGLYLEVALTHMET 121
QY 611 GAATTTCTC-----CATGCTCAACATACGGAAGAGTGGTGGCCCTT 655
DB 122 ASPTRPLEUVALARGASNVALTHRTYATGPRONISCUYSHISLECYSPHEINPROARG 141
QY 656 GGA-----GTCSTCTGAAGTTCTTAACCT-----TGCATCCTT 691
DB 142 GLYLEMETGLNRPHERHEGLNALANISPRORHPRORGERGLULYSYSERLYS 161
QY 692 TGGCTTCATGCTGACAGGAGGTCTAGATGAGCGCTTGATGAGAGGGG 742
DB 162 TRP-----LLEULEUGLUALSPRTYTRGLYSAGVALGILASNVAL 175
QY 743 -----GATGACATGCTGCTGCTGGGCTTGAGAGAAATAGAGATTTTACAGA 793
DB 176 THGLURPHEASPRSERLEU-----LEUATGAPNRPHEATHLEU 188
QY 794 CGCTATATACAGAAATGTTTGGAAATATAGCCAGACATATCAATGATCTTCCAA 853
DB 189 VALTHGLNHSIPROGLUVALILETYTRHASNGLINASNVALVALTRPSETLYSPHEGLU 208
QY 854 TCAATGGCTC----- 865
DB 209 THILERPHERHETHILSERGLYLEUHNISLYTALAPROVALPHEATGASPTYRVAL 228
QY 866 -----AAGGAATTCACAGATACCTGATCATGACGAGGAATACACATGAG 913
DB 229 PHEATGSRMETGLNGLURPHEITYRGLU-----ASPASNVALLEUITYRMEGLU 244

```

```

QY 914 ACTAGTACCAAGCTCTTAAGAAAGACCTGATCCCATTTGCCCTGTATGGCTTA 973
DB 245 ILEARGALALARGLEU-----PROVALTYR----- 253
QY 974 GAGGTATCTGATCTGCGAAGTGGAGAGCTGTCTTGAACCTGCTTGAATCGTG 1033
DB 254 GLNDEUSEGLYUHNISHISASRGLULTRPSEVALLYSHIRTYRGLNGLUVALALA 273
QY 1034 CAAGTTTCATCCAGGCGCCAAAGCCTACAGCCATTCATTAAGATGCAATGAAGAA 1093
DB 274 GLNLYSRPHEVALGLURHNIS--PROGLURPHEILEGLYLERLYLETLEYRSEAR 292
QY 1094 GCTGAGACACAGAAAGTTTCACTGTGTGACCTCTGTATGATCATCATGGGAT 1153
DB 293 HIS-ARGSERLYSASPVAL-----ALAVALLLEAGLUSERTILEARGME 307
QY 1154 CGCGAATGGCGAGCGC 1169
DB 307 TALAMETGLYLEUATRG 312

RESULT 10
US-09-310-363C-6
Sequence 6, Application US/09310363C
Patent No. 6388169
GENERAL INFORMATION:
APPLICANT: Mahajan, Pramod B.
APPLICANT: Shi, Jinru
APPLICANT: McELIVER, John
APPLICANT: Bowen, Benjamin
APPLICANT: Baszczyński, Christopher
TITLE OF INVENTION: Recl cDNAs and Uses Thereof
FILE REFERENCE: 0855
CURRENT APPLICATION NUMBER: US/09/310,363C
CURRENT FILING DATE: 1999-05-12
PRIOR APPLICATION NUMBER: 60/099,765
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/096,492
PRIOR FILING DATE: 1998-08-14
PRIOR APPLICATION NUMBER: 60/088,529
PRIOR FILING DATE: 1998-06-08
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 430
TYPE: PRT
ORGANISM: Zea mays
US-09-310-363C-6

Alignment Scores:
Pred. No.: 0.0974 Length: 430
Score: 96.50 Matches: 77
Percent Similarity: 33.85% Conservative: 53
Best Local Similarity: 20.05% Mismatches: 135
Query Match: 2.70% Indels: 119
DB: Gaps: 17

US-09-513-151-3 (1-2041) x US-09-310-363C-6 (1-430)

QY 32 CGAGCAGTCTCTGGGCGACT-----GGGCTCAGG 61
DB 97 ARGASPYALPROVALVALSERTHRGYSERPHEALALEUASPMETALALEUGLYTHRG 116
QY 62 GGCTGCAACAGCCCTACCTCTGTAGTATCTCGGGCCAGCAGGCGCAATGCC 121
DB 117 GLYLEUPROLYSGLY--ARGVALILEGLVALTYRGLYPROGLUALASERGLYSTR 135
QY 122 ACGCTGGCTTGCAAGCTA-----GGCCAGCGGCTCGCGGTGAGATGTCAGCGCT 172
DB 136 THLEUVALLEUHNISVALILEAGLUALGLYLSASNGLYLYTYRYSALAPHEVAL 155
QY 173 GACTCCATGCGAG-----GTCTATGAAGGCGCTAGACATCATCACCACCAAG 217
DB 156 ASPALAGLUNHISALALEUASPRIOALALEUAGLUSERTILEGLYVALSPRTHASN 175

```

```

QY 218 GTTTCGCCAAGAGCAAGAACTGCGCCGACACATGATGAGCTTTGGATCCTT 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 LeuLeuVal---SerGlnProAspGlyAlaGluGlnAlaLeuSerLeuValAspThrLeu 194
QY 278 GTGACCAATTACACAGTGTGACTTCAGAAATAGAGCACTGCTGTGATTGAGATATA 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 IleArgSerGlySerVal-----AspVal 202
QY 338 TTGGCCGAGCAAAATTCCTATTGTGTGGAGGACCAATTTATTCATTGATCTCG 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 203 ValValValAspSerValAlaAlaLeuVal----- 212
QY 398 CTCTGGAAGTCTTGTCATATACCAAGCCCGAGAGATGGCAGCTGAGAAAGTATTAC 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 ----- 212
QY 458 CGAAAGTGGAGCTTGAAAGAGAGATGGTCTTGACTTCACAAACGCCCTAAGCCAGTG 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 213 ProLysThrGluLeuAspLysLysLysLysLysLysLysLysLysLysLysLysLys 222
QY 518 GACCCAGAAATGCT-----GCCAAGCTGCATCCACATGACAAACGCAAGTGGCCAGG 571
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 223 AspAlaIleValAlaLeuGlnAlaLeuMetSerGlnAlaLeuArgLysLeuSerHis 242
QY 572 AGCTTGCAAGTTTGGAAAGAAAGCAATCTCATAGTGAATTTCTCCATGCTCAACT 631
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 SerLeuSerLeuSerGlnThrValLeuLeuPheIleAsnGlnIleArgAlaLysValAla 262
QY 632 ACGAAGAGTGTGTGTCCTT-----GGAGGTCCTGGAAGTCTCT 676
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 ThrPheGlyPheGlyGlyProThrGluValIleThrSerGlyGlyAsnAlaLeuLysPhe 281
QY 677 AACCTTGACATCTTGCTGCTCATGCTGACAGGCACTTCTAGATGAGCGCTTG----- 730
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 -----TyrAlaSerValArgLeuAsnIleArgArgIleGlyPhe 294
QY 731 -----GATAGAGGCGATGACATGCTTGCTGCTGCGCTCTTGAG----- 772
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 LeuLysGluGluGluGluThrIleGlySerGlnValAlaValLysIleValLysAsnLys 314
QY 773 -----GACCTAAGAGATTTTTCACAGACGCTATATACAGAAATGTTTGC 817
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 HisAlaProProPheLysThrAlaGlnPheGluLeuGluPheGly---LysGlyIleCys 333
QY 818 GAAATAGCCAGACTATACATGATGATC----- 847
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 ArgSerSerGluLeuPheGluLeuGlyLysHisLysLysLeuIleGlnLysThrGlyGly 353
QY 848 -----TCCCAATCAATGCTGCTCAG-----GAATTCACAGATCCTG 886
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 AlaTyrTyrArgPheAsnAspMetSerPheLysGlyLysAsnAsnLeuLysSerTyrLeu 373
QY 887 ATCACTAGAGAAATGACACATGAGACTAGTAACACACTTCTAAGAAAGAGCTGCT 946
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 ---ThrGluAsnLysSerValAlaAsnAspLeuGluThrLysLeuArgArgLeuMetGly 392
QY 947 CCCATTGCCCCCTGCTCATGCTTAAAGGTATCTGATGTCCTCCAGAGTGGAGAGCTCT 1006
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 393 ThrGluAlaProLysGluGlnGlnAlaGlnAspSerSerProSerAspLeuProGluGln 412
QY 1007 GTTCTTGAACT 1018
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 413 ValValThrPro 416

```

```

RESULT 11
US-08-557-122A-26
; Sequence 26, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: HJORT, Carsten Mølland
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5879664 disk of No. 5879664th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,122A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-9655
TELEFAX: 212-878-0123
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3052 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-557-122A-26

Alignment Scores:
Pred. No.: 0.465 Length: 3052
Score: 95.50 Matches: 86
Percent Similarity: 36.91% Conservative: 62
Best Local Similarity: 21.45% Mismatches: 177
Query Match: 2.67% Indels: 76
DB: Gaps: 16

US-09-513-151-3 (1-2041) x US-08-557-122A-26 (1-3052)
QY 10 GATGCGCTGCGGCGCGGACAGACAGTCTGCGGAGGCGCTCAGGCGCTGCA 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1028 AspAspIleProPheGlyLeuThrAlaSerSerAspAlaIleAlaIleGluThrLeu 1047
QY 70 ACGGACCTACTCTTGATGATTCCTCGGGCCACGCGGACCGGCAATCCAGCTGCG 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1048 ValAspSerSerGluValValIleGlyPhePheLysAspValThrSerAspAlaIle 1067
QY 130 GTTGCAGCTAGGCGCGGCTGCGGCTGAGATCGTC----- 166
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1067 AlaYsgLysPhe-----LeuLeuAlaIleGluSerValAspAspIleProPheGlyIle 1085
QY 167 -----AGCGTGATCCATCAGGTCTATGAGCGCTAGACATCACCACCAAGTTTC 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1085 rSerSerAlaAspAspLeuProAlaTyrLeuAlaAsnGluThrPheValIleThrProValIle 1105
QY 223 TGCCCAAGACAGAGATCTGCCGCGCACCAATGATCAGCTTGTGTGATCTCTGTGAC 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1105 eValGlnSerGlyLysIleAlaAspAlaAspPheAsnAlaThrPheTyrSerMetAlaAsnLys 1125
QY 283 CAATTACACAGTGTGAGATTGAGAAATAGACACACTGCTCTGATTTGAAGATATATTTC 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 sHisPheAsnAspTyrAspPheValSerAlaGlnAsnAla-----AspValPhe 1142
QY 343 CCGA-----GACAAATTCCTATTGTTG-----GGAGAAC 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1142 rLysTyrGlnLeuAspLysAspGlyValValLeuPheLysLysPheAsnGluGlyArgAs 1162
QY 376 CAATTATTACATTGATCTCTGCTGTAAGTCTTGTCAATTACCAAGCCCGAGAGAT 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1162 nAsnPheGluGluIleThrLysGluLysLeuLeuAspPheIleLysHisAsnGlnIle 1182

```

```

Y 436 GGGAGCTGCAAAAGTATTCACCGAAAAGCTTGAAAAGGAAGGTGGCTGTACT 495
Db      :          :          |||          :|||::|||::|||::|||::|||
Y 1182 uproleu---ValilevalPheSerIygtgInleuAsprlys---AsrglyvalValle 1200
Y 496 TCACAAACGCSTT-----AGCCAGCTGAGCACCAGAATGGC 531
Db      :   |::|::|          :|||::|||::|||::|||::|||
Y 1200 upheliysLyPhasPrcluglYArgasnphhegluglYglValtIthrlYsglLYe 1220
Y 532 TGGCAAGCTGCATCCACATGACAAACGAAAGTGGCCAGAGCTTGCAAGTTTGAAGA 591
Db      :   |::|::|          :|||::|||::|||::|||::|||
Y 1220 uLeuArPheIleLyuShIsAsnGlInleuProLeu-----ValilevalPheSerly 1237
Y 592 AACAGGAATGTCSCATAGTGAATTGTCATGTCGTAACAATACGAGAAAGAGCGTGCC 651
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1237 sTygInleuAsrLyshsrplYvalValleuPheLyshrsPhesrglUglYArgsnAs 1257
Y 652 CCTTGAGCTGCTGTGAAGTCTCTAAC-----CC 681
Db      :   |||   :   :   :   :   :   :   :   :   :   :   :   :
Y 1257 nPheglUGlYglValtIthrlYsglnshInleuAerPheIleLyshIsAsnGlInleuPr 1277
Y 682 TTGCATGCTT-----TGGCTCATGCTGACCCAGCAGTTCTAGATGAGCC 726
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1277 oLeuValtIlevalPheSerIrtgtgInvalInlsGlnAsrglYvalValleuPheLy 1297
Y 727 CTTGATAAGAGGGTGGATGACATGCTGCTGCTGGGCTCTTGGAAGAACTAAGAGATT 786
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1297 srPhesrglUGlYArgasnphhegluglYglValtIthrlYsglLYeLeuAsrPh 1317
Y 787 T-----CACAGACGCTATATACAGAAGATGTTTCGAAAAATAGCCAGACTATACAA 840
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1317 eIleLyshIsAsnGlInleuProLeuValtIlevalPheSerlyTygInleuSerGlnAs 1337
Y 841 TGGAATGTCCTCAATGCTGCTTCAAGAAATTCACGAGTACGATGATACT---GAGGG 897
Db      :   |||::|          :|||::|||::|||::|||::|||
Y 1337 rglYval-----ValleuPheLyshrsPhesrglUGlYArgasnphheglUGl 1354
Y 898 AAAATGACACACTGGAGACTAGTAACACAGCTTCAAAGAAAGAACTGGTCCCATGTGC-- 955
Db      :   |   |   :   :   :   :   :   :   :   :   :   :   :   :
Y 1354 yAsrPheIthrlYAsrpsnshleuAsnphelIelysSeraNgInleuProLeuValtl 1374
Y 956 -----CCCCGTGTATGG 969
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1374 eaSPAsrPheLyshLeuSerIleTyrtleuProseralameLasrglProvaltItyAs 1394
Y 970 CTTAGAGATACATGCATGCTCGAAGTGGGAGAGCTGTGTTGAAACCTGCTTGAAT 1029
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1394 nglyLyshAlaAspIleAla-----AspAlaAsrvalPheglULystrPlueInva 1412
Y 1030 C 1030
Db      :   :   :   :   :   :   :   :   :   :   :   :   :   :   :
Y 1412 l 1412

RESULT 12
US-09-262-666-26
; Sequence 26, Application US/09262666
; Patent No. 6346244
GENERAL INFORMATION:
APPLICANT: Hjort, Carsten Molland
TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ad No. 6346244O No. 6346244dlisk of No. 6346244th America, Inc
STREET: 405 Lexington Avenue, 64th Floor
City: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

```

1 CURRENT APPLICATION DATA:  
 2 APPLICATION NUMBER: US/09/262,666  
 3 FILING DATE:  
 4 CLASSIFICATION:  
 5 PRIOR APPLICATION DATA:  
 6 APPLICATION NUMBER: US 08/557,122  
 7 FILING DATE: 11-DEC-1995  
 8 ATTORNEY/AGENT INFORMATION:  
 9 NAME: Lambiris, Elias J.  
 10 REGISTRATION NUMBER: 33,728  
 11 REFERENCE/DOCKET NUMBER: 3980.204-US  
 12 TELECOMMUNICATION INFORMATION:  
 13 TELEPHONE: 212-867-0123  
 14 TELEFAX: 212-878-9655  
 15 INFORMATION FOR SEQ ID NO: 26:  
 16 SEQUENCE CHARACTERISTICS:  
 17 LENGTH: 3052 amino acids  
 18 TYPE: amino acid  
 19 STRANDEDNESS: single  
 20 TOPOLOGY: linear  
 21 MOLECULE TYPE: peptide  
 22 US-09-262-666-26

Alignment Scores:

Pred. No.:	0.465	Length:	86
Score:	95.50	Matches:	86
Percent Similarity:	36.91%	Conservative:	62
Best Local Similarity:	2.14%	Mismatches:	177
Query Match:	2.67%	Indels:	76
DB:	4	Gaps:	16

US-09-513-151-3 (1-2041) x US-09-262-666-26 (1-3052)

OY	10	GATGGCTCCGTGGCGGCTGCACAGACAGTCTCTGGGCAATGGGCTCAGGGGCTCA	69
		:        :        :        :        :        :        :	
Db	1028	AspAspIleProPheGlyLeuThrIlaSerSerAspSalaIaLaIaIaIaIuThrLeu	1047
OY	70	ACGAGCCCTACCTCTTACTGATTCGCGGGCCACGGGACCCGGCAATCCAGCTGC	129
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1048	ValAspSerSerGlu-ValValIaIleGlyPhePheIlyAspValThrSerAspAlaI	1067
OY	130	GTTCCAGCTAGCGCCAGCGGCTCGGCGGTGAGATCGTC	166
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1067	AllySGluPhe-----LeuLeuIaIaIaGluSerValaAspAspIleProPheGlyIleSe	1085
OY	167	----AGCGCTGACTCCATGCAGAGCTCTATGAAGCCCTAGACATCATCACCACAGCTTC	222
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1085	rSerSerIlaAspSerpLeuProAlaIyThrLeuIaIaAsnGluThrPheValThrProValI	1105
OY	223	TGCCACAGACAGAGATCTGCGCGGACCAATGATCAGCTTGTGGATCCCTTGTCAC	282
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1105	eValGlnSerGlyIySlyIleAspIaAspPheAsnAlaThrPheTySerMetAlaAsnIy	1125
OY	283	CAATTACACAGCTGTGCATCTCCAGAAATACGACCACTGCTCGATTGAATATATTCG	342
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1125	SHisPheAsnAspTyIyrAspPheValSerIaIaGluAsnIaIa-----AspValPheSe	1142
OY	343	CCGA-----GACAAATTCCTATTTGTTG-	375
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1142	rLyTyGlnLeuAsnSprIyAspGlyValValIleuPheIySlySPheAspGluGlyIaYArg	1162
OY	376	CAATTATTACTTGATCTCTGCTCTGGGAAAGTTCTTGTCATTACCAAGCCCCAGGAAT	435
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1162	nasnPhelGluIyIuIleThrIySglIyIleuLeuAspPheIleTySHisAsnGlnIle	1182
OY	436	GGGCGACGAGAAATGATTGACCGGAAAGTGAGCTTGAAGAGAGATGGCTGTACT	495
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1182	uProLeu--ValIleValPheSerIyTyGlnIleuAspIyS--AspGlyValValIle	1200
OY	496	TCACAAAGCGCTA-----ACCCAGGTGAGACCCCAAGATATGC	531
		:    :    :    :    :    :    :    :    :    :    :    :	
Db	1200	uPheIyIySPheAspGluIyIaYArgAsnAsnPhelGluIyGluValThrIySglIyIySe	1220

```

QY 532 TGCGAAGCTCATCATGACAAAGCAAGTGGCCAGAGCTTGCACTTTTGAGA 591
Db 1220 ulcuaspphelelyshisnsglnleuprou-----Vallelepheserly 1237
QY 592 AACGGAATCTCATAGTGAATTTCTCCATCGTCAACATACGAGAGAGTGGTGC 651
Db 1237 styrnglnleuasphelelyspolylvalleuPhelyslyspheaspclncllyragsnas 1257
QY 652 CCTGGAGGCTCTGAGAGTCTCTAAC-----CC 661
Db 1257 nphcgluglyglvalthrllysnglnleuasphelelyshisnsglnleupr 1277
QY 682 TTGCATCTT-----TGGCTTCATGCTGACGAGAGCTTGATAGAGCG 726
Db 1277 oleuvallelevalpheserlytyrlyglvalhlsnglnaspcllyvalleuPhelysly 1297
QY 727 CTTCGATTAAGAGGTGATGATGCTTGGCTGCTGCTGCTGAGAGTAAAGATTT 786
Db 1297 spheaspgluglylYragsnasnphcgluglyglvalthrllysnglnleuasphe 1317
QY 787 T-----CACAGACGCTATATACAGAAGATTTTCGGAATAACGACATATCAACA 840
Db 1317 eilelyshisnsglnleuprouleuVallelevalpheserlytyrlyglleuSerlgnas 1337
QY 841 TGCTATCTTCATCAATGCTTGAAGATTTTCAGAGTACCTGATCACT---GAGGG 897
Db 1337 pgllyval-----ValleuphelelyshisnsglnleuprouleuVallelevalpheserly 1354
QY 898 AAAATGCACATGAGAGCTAGTAAAGAGCTTCTAAAGAAAGAGCTGGTCCATTGTC- 955
Db 1354 yaspleuThrlYaspsnleuasnphellelyserasnsglnleuprouleuValle 1374
QY 956 -----CCCCCTGCTATGC 969
Db 1374 easpasphelelyseuSerlelyrleuprouSerlgnasleuasphelelyshisnsglnleuprou 1394
QY 970 CTTCAGAGTATGATGCTGAGAGTGGAGAGCTGCTCTTGAACCTCTCTGGAAT 1029
Db 1394 nglyshlysalasphelelele-----AspalaasphelelyshisnsglnleuprouleuValle 1412
QY 1030 C 1030
Db 1412 1 1412

RESULT 13
US-08-836-442-3
Sequence 3, Application US/08836442
Patent No. 5990293
GENERAL INFORMATION:
APPLICANT: DOCHERTY, Andrew, J.P.
APPLICANT: SLOCUMBE, Patrick, M.
TITLE OF INVENTION: A HUMAN METALLOPROTEINASE
TITLE OF INVENTION: VARIANTS THEREOF AND DNA SEQUENCES CODING THEREFOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,442
FILING DATE: 01-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/02181
FILING DATE: 13-MAR-1997

```

```

APPLICATION NUMBER: GB 9612150.4
FILING DATE: 11-JUN-1996
APPLICATION NUMBER: GB 9526229.1
FILING DATE: 21-DEC-1995
APPLICATION NUMBER: GB 9521498.7
FILING DATE: 20-OCT-1995
APPLICATION NUMBER: GB 95521495.3
FILING DATE: 20-OCT-1995
APPLICATION NUMBER: GB 9518023.8
FILING DATE: 05-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47425
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 529 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 5990293e
US-08-836-442-3

Alignment Scores:
Pred. No.: 0.213 Length: 529
Score: 94.00 Matches: 72
Percent Similarity: 32.93% Conservative: 63
Best Local Similarity: 17.56% Mismatches: 161
Query Match: 2.638 Indels: 114
DB: 2 Gaps: 19

US-09-513-151-3 (1-2041) x US-08-836-442-3 (1-529)
QY 413 GTCAATACCAAGCCGAGAGTGGCGACATGAGAAAGTATGACGGAAGTGAAGCTT 472
Db 95 lilehtrhserProglinlelel----- 102
QY 473 GAAAGAGAGAT-----GGTCTTGTACTTCACAAAGCCCTAACGAGTGGAC 520
Db 103 -----Aspaspyslytyrlyglnglyhlsleuasnsglnlysalaser 120
QY 521 CCAGAAATGGCTGCCAAGCTGCATCCACAT-----GACAAGCCAAAGTGGCC 568
Db 121 lleserthrcysarglyleuarglytyrPheserlgnlyasphelelyshisnsglnleuprouleuValle 140
QY 569 AGGAGCTGCAAGTTTGAAGAAAGCAAGTCTCATAGTGAATTTCTCCATGTCAC 628
Db 141 gluprouleuSerlgnlelyshisnsglnleuprouleuVallelevalpheserlytyrlyshisnsglnleuprouleuValle 160
QY 629 CATACGGAAGAGTGGTGGTCCCTGGAGTCTCTGAAAGTCTCTAACCTTGC-- 685
Db 161 Aspglulys-----Asnlyraserlthrcysgly 170
QY 686 -----ATCCTTGGCTTCATGCTGACGACGACGACTTCAATGAGCCCTGGATAG 736
Db 171 Metaspelyalleartrahls-----Aspleu 180
QY 737 AGGAGTGAATGATGCTGCTGCTGGCTTGGAGAACTAAGAT----- 784
Db 181 glnglnsnlaleuProhlatnlyshisnsglnleuprouleuVallelevalpheserlytyrlyshisnsglnleuprouleuValle 200
QY 785 TTTCAGAGAGCTATATACAG-----AAGAAATGTT 814
Db 201 glnglnsglnlyshisnsglnleuprouleuVallelevalpheserlytyrlyshisnsglnleuprouleuValle 220
QY 815 TCGGAATATGACGAGCTATCAACATGATCTTCCA-----TCAAT 859
Db 221 AsnglnasnsglnasphelelyshisnsglnleuprouleuVallelevalpheserlytyrlyshisnsglnleuprouleuValle 240

```

```

QY 860 GGCCTCAAGATTCACAGTACCTGATCACTAGAGAAATGACACTGAGACTAGT 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 LeuTyrLysLysLeuAsnThrHisValAlaLeuValGly---MetLuletrPrThrAsp 259
QY 920 AACGACCTTCAAGAAAGAGACCTGCTCCATTGCCCCCTGCTCATAGGCTTAGAGTA 979
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 LysAspLysLleLys-----LleThrProAsnLaseSerPheThrLeu 273
QY 980 TGTGATGTCGAAAGGAGAGAGTGTCTTTGAA----- 1015
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 GluAsnPheSerLysThrArgLysValLeuSerArgArgLysArgHisAspIleAla 293
QY 1016 -----CTGCTCTTCAAAATCGTCAAGT----- 1039
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 GluLeuLleThrAlaThrGluLeuAlaGlyThrValGlyLeuAlaPheMetSerThr 313
QY 1040 -----TTTCATCAGAGGCCACAGACCTGACAGCCATCCAAATA 1075
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 314 MetCysSerProTyrSerValGlyValAlaGlnAspHisSerAspAsnLeuLeuArgVal 333
QY 1076 -----AGATCCATACATGATGAGCTGAAACAGACAGATTATCAC-----CTG 1120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 AlaGlyThrMetAlaHisGluMetGlyHisAsnPheGlyMetPheHisAspAspTyrSer 353
QY 1121 TGTGACCTGTGATGATCATCATTTGGGATTCGGGATTCGGCAGCCGACATAAATCC 1180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 CysLysCysProSerThrLleCysValMetAspLysAlaLeuSerPheTyrLleProThr 373
QY 1181 AAA---TCCCACTTGAAACCACTGAAAGAAAGAAAGATTGACATGCTGCTCAAC 1237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 374 AspPheSerSerCysSerArgLeuSerTyrAspLysPhePheGluAspLysLeuSerAsn 393
QY 1238 ACCATGAAAGTACAGATGTTTCCCACTATACAAAGAAACCTTAAGGGAAGGATCC 1297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 CysLeuPheAsnAlaProLeuProThrAspLleLleSerThrProLecysGlyAsnGln 413
QY 1298 CCAGGGCAGATGATCAAGCTGAATGACAGGTTTAGACATCTCCAGT----- 1351
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 LeuValGluMetGlyGluAspCysAspCysGlyThrSerGlyThrCysLysIleLysAla 433
QY 1352 GCCTTTGAAAGGTGGGG-----ATCCAGTTTCAGAGGAGGAGGTAT 1396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 434 ThrPheGlnCysAlaLeuLleGlyCysGlyLysGlnPheLysAlaGlyMet 453
QY 1397 GTTGTCTCCAGTCTGGCCAAAGAGTGC 1426
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 ValCysArgProAla---LysAspGluCys 462

```

## RESULT 14

```

US-09-040-725A-1
; Sequence 1, Application US/09040725A
; Patent No. 6399584
; GENERAL INFORMATION:
; APPLICANT: Institut Curie
; APPLICANT: CNRS
; APPLICANT: Arpin, Montique
; APPLICANT: Crepaldi, Tiziana
; APPLICANT: Gautreau, Alexis
; APPLICANT: Louvard, Daniel
; TITLE OF INVENTION: Pharmaceutical composition containing ezrin mutated
; FILE REFERENCE: 39108200100
; CURRENT APPLICATION NUMBER: US/09/040/725A
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 586
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-040-725A-1

```

Alignment Scores:

```

Pred. No.: 0.228
Score: 94.00
Percent Similarity: 32.068
Best Local Similarity: 19.148
Query Match: 2.63%
DB: 4
Gaps: 16

US-09-513-151-3 (1-2041) x US-09-040-725A-1 (1-586)

QY 164 GTCAGGCGTACCTCATGAGCTATGAAAGGCTAGACATCAACAAAGTTTCT 223
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 ValAspAlaLeuGlyLeuAsnLleTyrGluLysAspAspLysLeuThrProLysIleGly 239
QY 224 GCCCAAGACACAGAAATCTGCCGACACATGATGACTT----- 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 PheProTyrSerGluLle-----ArgAsnLleSerPheAsnAspLysPheVal 256
QY 266 -----GTGATCCTCTTGTGACCAATTCACAGTGTGACTTCAAAATAGACA 316
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 257 IleLysProLleAspLysLysAlaProAspPheValPheTyrAlaProArgLeuArgIle 276
QY 317 ACTGCTGTGAT-----GAAGATATATTGCCCCGAGCAAAATT 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 AsnLysArgLleLeuGlnLeuCysMetGlyAsnHisGluLeuTyrMetArgArgArgLys 296
QY 356 CCTATTGTTGTGGGAGAACCAATTATTTCATTGAATCTGCTGCTGAAAGTTCTTTC 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 ProAspThrLleGluValGlnGlnMetLysAlaGlnAlaArgGluLysHisGlnLys 316
QY 416 AATACCAAGCCCGAGAGATGGGCACTGAGAAAGTATTGACCGAAATGGAGCTTGA 475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 317 GlnLeuGluArgGlnGlnLeuGlnLysLysArgArgLysThrValGluArgGlu 336
QY 476 AAGGAGATGCTTCTTACTTCACAAACCTTAAGCCAGTGGAGCCCAAAATGCTGCC 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 LysGlu-----GlnMetMetArgLysGluGluLeuMetLeu 349
QY 536 AAGCTCATCCATGACACAAAGC-----AAAGTG 565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 ArgLeuGlnAspTyrGluGluLysThrLysLysAlaGluArgLysLeuSerGluGlnIle 369
QY 566 GCCAGAGCGTGCAGTTTGTGAAGA----- 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GlnArgAlaLeuGlnLeuGluGluArgLysArgAlaGlnGluAlaGluArgLys 389
QY 593 -----ACAGGAATCTCTCATAGTAATTTCTCCATGCTCAACATACGAA 637
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 GluAlaAspArgMetAlaAlaLeuArgAlaLysGluGluLeuGluArgGln 406
QY 638 GAAGTGTGTCTCCCTTGAGGCTCTTGAACTCTTAACCTTGATCTTGGCTT 697
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 ----- 406
QY 698 CATGCTGACCGCATCTTATAGATGAGCGCTTGATGAAGGCTGATGATGCTTGTCT 757
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 AlaValAspGlnLleLysSerGlnGlnLeuAlaAlaGluLeuAlaGluTyrThrAla 426
QY 758 ---GCTGGCTCTTGGAGAACTAGAGATTTTTCACAGCGCTTAATCAGAAATGTT 814
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 LysLleAlaLeuLeuGluGluAlaArg-----ArgArgLys 438
QY 815 TCGGAAATAGCCAGACATCATCAATGTTATCTTCAATCAATTTGGCTTCAAGAAATT 874
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 GluAspGluValGlnLleTyrPheGlnHisArgAla-----LysGluAla 452
QY 875 CAGGACTACCTGATCACTAGAGGAAATGACACACTGAGACTAGTAACAG-----CTT 928
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 453 GluAspAspLeuVal-----LysThrLysGluLleLeuHisLysLeu 465
QY 929 CTAAGAAAGAGACCTGCTCCATGTCCTCTATGAGCTTAGAGGTATCTGATGTC 988
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 ValMetThrAlaProProProProProProProProValTyr----- 478

```

QY 989 TCGAGTGGGAGAGACTGTTCTTGAACCTGCTTGAATGCGCAAGTTTCATCCAG 1048  
 Db 479 -----GluProValSerThrHisValGlnSerLeuGln 490  
 QY 1049 -----GGCCACAGCCTACAGCCACTCCATTAAGATGCCATACATGAAGCTGAGAAC 1102  
 Db 491 AspGluGlyAlaGluProThrGly----- 498  
 QY 1103 AAGAGAGTTATCACCCTGTGACCTCTGTGATCGAATCATCATTTGGGATCGCAATGG 1162  
 Db 499 -----TyrSerAlaGluLeuSerSerGluGlyIleArgAspAspArg----- 512  
 QY 1163 GCAGGCGCATTAATCCAAATCCCACTTGAAACCACTGAGAAAGAAAGAAAGATGGAC 1222  
 Db 513 -----AsnGluGluValArgIleThrGluAlaGluValAsnGluArgValGln 528  
 QY 1223 TCAGATGCTGCACACACATAGAAAGTCAAGCTGTTCCCGCATATACAA 1276  
 Db 529 ArgGlnLeuValThrLeuSerSerGluLeuSerGlnAlaArgAspGluAsnLys 546

## RESULT 15

US-08-836-943-2  
 ; Sequence 2, Application US/08836943  
 ; Patent No. 5965391

## GENERAL INFORMATION:

APPLICANT: Reinscheid, Dieter  
 APPLICANT: Elkmanns, Bernhard  
 APPLICANT: Sahm, Hermann  
 TITLE OF INVENTION: DNA WHICH REGULATES GENE EXPRESSION IN  
 NUMBER OF SEQUENCES: 3  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: The Firm of Karl F. Ross, PC  
 STREET: 5676 Riverdale Ave.  
 CITY: Bronx  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10471

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/836,943  
 FILING DATE: 08-MAY-1997  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Myers, Jonathan  
 REGISTRATION NUMBER: 26,963  
 REFERENCE/DOCKET NUMBER: 20357  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (718) 884-6600  
 TELEFAX: 718/601-1099  
 TELEX: 620428

## INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 739 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-836-943-2

## Alignment Scores:

Pred. No.: 0.266 Length: 739  
 Score: 94.00 Matches: 57  
 Percent Similarity: 37.24% Conservative: 70  
 Best Local Similarity: 20.53% Mismatches: 136  
 Query Match: 2.63% Indels: 78  
 DB: 2 Gaps: 15

US-09-513-151-3 (1-2041) x US-08-836-943-2 (1-739)

QY 89 GTGATTCGCGGGCCAGCGGCACC-----GGCAATCCACGCTGGCGTTG 133  
 Db 336 ValTyrIleGlyArgAsnGlyThrGluLeuValLeuHisGlyArgSerLeuPheVal 355  
 QY 134 CAG---CTAGGCGCAGCGGCTGGCGGTAGATTCGACGCGGTGATCCATGAGGTGAT 190  
 Db 336 ArgAsnValGlyHisLeuMetGlnAsnProSerIleLeuAspGlyGluIlePhe 375  
 QY 191 GAAGGC---CTAGCATCATCCACCAACAGGTTTCTGCC----- 226  
 Db 376 GluGlyIleMetAspAlaValLeuThrThrValCysAlaIleProGlyIleAlaProGln 395  
 QY 227 CAAGACGAGAAATATGCGCGCACCATGATCAGCTTTGGATCCTTGTGACCAAT 286  
 Db 396 AsnLysMetArgAsnSerArgLysGlySerIleValLysProLysGlnHisGly 415  
 QY 287 TACACAGTGGGAGCTTACAGAAATGACCAACTGCTGATGTAAGATATATTGCGCGA 346  
 Db 416 ProGluGluValAlaPheThrAsnGluLeuPheGlyArgValGlnAspLeuAspLeu 435  
 QY 347 GACAAATTCCTATTGTTGTGGGA-----GGAACCAATTATTACATT 388  
 Db 436 ProArgHisThrLeuLysValGlyValMetAspGluArgThrSerValAsnLeu 455  
 QY 389 GAATCTGCTCTGGAAGTT-----CTTGCAATACCAAGCCCGAGAG 433  
 Db 456 AspAlaSerIleMetGlnValAlaAspArgLeuAlaPheIleAsnThr----- 471  
 QY 454 ATGGCACTGAGAAAGTATGACCGA-----AAGAGGACCTTGAAAGAGAGAT 484  
 Db 472 -----GlyPheLeuAspArgThrGlyAspGluIleHisThrSerMetGluAla 487  
 QY 485 GGTCTTGACTTACACAAAGCGCTAGCCAGGTGACCCAGAAATGGTCCAGAGCTGAT 544  
 Db 488 GlyAlaMetValArgLysAlaAspMetGlnThrAlaProTyrGlnAlaTyrGlnAsn 507  
 QY 545 CCACATGACAAACGCAAGGTGGCCAGAGCTTGCAAGTTTGAAGAACAAGAACTCTCT 604  
 Db 508 AsnAsnValAspAlaGlyIleGlnArgGlyLeuProGlyLysAlaGlnIleLysGly 527  
 QY 605 -----CATAGTAATTTCTCATCGCAACATACGAAAGAAAGT 643  
 Db 528 MetTrpAlaMetThrGluLeuMetAlaGluMetLeuGluLysLys----- 542  
 QY 644 GGTGTCCTCCCTGGAGGTCCTCTGAAGTTCTCTAAC-----CCTTGC 685  
 Db 543 -----IleGlyGlnProArgGluGluAlaAsnThrAlaTrpValProSerProThr 559  
 QY 686 ATCCTTGGCTTCATCGACGACGACGACGATCTAGATGAGCGCTTGGAATGAAGGGTGAT 745  
 Db 560 GlyAlaThrLeuHisAlaThrHisTyrHisLeuValAsp---ValPheLysValGlnAsp 578  
 QY 746 GACATGCTGCTGCTGGCGCTCTTGAGGAGACTAGAA----- 781  
 Db 579 GluLeuArgAlaAlaGlyArgArgAspSerLeuArgAsnIleLeuThrIleProThrAla 598  
 QY 782 -----GATTTTCACAGACGCTATATATCAGAAAGATTTTTCGAAATAGCAGAGAC 832  
 Db 599 ProAsnThrAsnTrpSerGluGluGluLysGluLysGluMetAspAsnAsnCysGlnSer 618  
 QY 833 -----TATCAACATGGTATCTTCCATCAATGATGGCTTCAAG 868  
 Db 619 IleLeuGlyTyrValValArgTrpValGlnHisGlyValGlyCysSer---LysValPro 637  
 QY 869 GAATTTCAGCAGTACCTGATCATCTAGGAGGAATGCCAGACGAGACCTAGTACCAAGTT 928  
 Db 638 AspIleHisAspIleAspLeuMetGlnAspArgAlaThrLeuArgIleSerSerGlnMet 657  
 QY 929 CTA 931  
 Db 658 Leu 658

Tue Apr 22 09:11:09 2003

us-09-513-151-3.n2p.rai

Page 15

Search completed: April 21, 2003, 18:52:17  
Job time : 42.0286 secs

---

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus.n2p model

Run on: April 21, 2003, 17:14:14 ; Search time 24.423 Seconds  
(without alignments)  
6932.240 Million cell updates/sec

Title: US-09-513-151-3  
Perfect score: 3575  
Sequence: 1 CTGCCATTAAGATGCGTCG.....TTTACAGAGAAAAAAAAAAAA 2041

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+.n2p.model -DEV=xlh  
-O=/cgn2.1/USPPO.spool/US09513151/runat.15042003.141143.26366/app.query.fasta.1.2446  
-DB=SwissProt\_40 -QFMT=fasta -SUFFIX=72p.rsp -MINMATCH=0.1 -LOOPT=0  
-LOOPT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human.0.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09513151.cgn.1.1.25 @runat.15042003.141143.26366 -ICPU=3  
-NO\_XIPXY -NO\_MMAP -LARGEOVER -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	548	15.3	428	1 MODS_YEAST	P07884 saccharomyc
2	417	11.7	314	1 MIAA_BACHD	O9kac3 bacillus ha
3	403.5	11.3	314	1 MIAA_BACSU	O81795 bacillus su
4	397.5	11.1	305	1 MIAA_LISMO	O8y713 listeria mo
5	395.5	11.1	305	1 MIAA_LISIN	O92c59 listeria in
6	378.5	10.6	309	1 MIAA_CLOAB	O97121 clostridium
7	377.5	10.6	310	1 MIAA_CLOPE	O8x185 clostridium
8	363	10.2	323	1 MIAA_RALSO	O8xw60 ralsconia s
9	349	9.8	322	1 MIAA_PSEPU	O30762 pseudomonas
10	348	9.7	305	1 MIAA_THEMA	O9wv25 thermotoga
11	343.5	9.6	314	1 MIAA_CHETR	O84771 chlamydia t
12	342.5	9.6	316	1 MIAA_PASMU	O9cm57 pasteurilla
13	339	9.5	311	1 MIAA_STEAM	O94495 haemophilus
14	334.5	9.4	311	1 MIAA_HAEIN	O9kvi2 vibrio chol
15	334	9.3	315	1 MIAA_VIBCH	O9chn2 lactococcus
16	333	9.2	294	1 MIAA_LACIA	O8ylm2 anababena sp
17	329.5	9.2	295	1 MIAA_AKNSP	O9hul9 pseudomonas
18	328.5	9.2	323	1 MIAA_PSEAE	

19	327.5	9.2	314	1 MIAA_CHLNU	O9p1f7 chlamydia m
20	324.5	9.1	311	1 MIAA_MYCLE	P46811 mycobacteri
21	324.5	9.1	314	1 MIAA_MYCTU	O33232 mycobacteri
22	321.5	9.0	313	1 MIAA_NEIMA	O9juu5 neisseria m
23	320	9.0	294	1 MIAA_STEPM	O971w5 streptococc
24	319.5	8.9	342	1 MIAA_CHLPPN	O9z6z6 chlamydia p
25	318.5	8.9	299	1 MIAA_STEPM	O9a059 streptococc
26	318.5	8.9	313	1 MIAA_STEPM	O8ziw3 yersinia pe
27	314	8.8	305	1 MIAA_AQUAE	O67162 aquifex ae
28	312.5	8.7	317	1 MIAA_XYLEA	O9p156 xyella fas
29	311	8.7	316	1 MIAA_ECOLI	P16384 escherichia
30	310	8.7	312	1 MIAA_ECO57	O8xnd3 escherichia
31	305	8.5	312	1 MIAA_STEPM	O69967 streptococ
32	302	8.4	316	1 MIAA_SALTI	O8z186 salmone
33	301.5	8.4	316	1 MIAA_BORBU	O51761 borrelia bu
34	301	8.4	316	1 MIAA_SALTY	P37724 salmone
35	295.5	8.3	312	1 MIAA_HELPI	O9z1j7 helicobacte
36	288	8.1	303	1 MIAA_RHIME	O92nr2 rhizobium m
37	287.5	8.0	311	1 MIAA_HELPI	O25961 helicobacte
38	286.5	8.0	306	1 MIAA_DEIRA	O9rtt6 delnoco
39	281.5	7.9	326	1 MIAA_RICCN	O92hw4 rickettsia
40	280	7.8	313	1 MIAA_STY3	P74040 synecocyst
41	278.5	7.8	303	1 MIAA_NEMIB	O9jzr0 neisseria m
42	272	7.6	298	1 MIAA_AGR5	P38436 agrobacteri
43	265.5	7.5	321	1 MIAA_RHILU	O98k14 rhizobium l
44	259	7.2	315	1 MIAA_BUCAI	P57632 buchnera ap
45	256	7.2	295	1 MIAA_CAUCR	O9a6j5 caulobacter

## ALIGNMENTS

RESULT 1  
MODS\_YEAST  
ID MODS\_YEAST STANDARD; PRT; 428 AA.  
AC P07884; Q12203;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE tRNA isopentenyltransferase (EC 2.5.1.8) (Isopentenyl-diphosphate:  
DE tRNA isopentenyltransferase) (IPTase) (IPTP).  
GN MOD5 OR YOR274W.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=6712703; PubMed=3031457;  
RA Najarian D., Dhanich M.E., Martin N.C., Hopper A.K.;  
RT "DNA sequence and transcript mapping of MOD5: features of the 5'  
RT region which suggest two translational starts.";  
RL Mol. Cell. Biol. 7:185-191(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=5288c;  
RX MEDLINE=97051594; PubMed=8896271;  
RA Cheret G., Bernardi A., Sor F.J.;  
RT "DNA sequence analysis of the VP1-SNF2 region on chromosome XV of  
RT Saccharomyces cerevisiae.";  
RL Yeast 12:1059-1064(1996).  
RN [3]  
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.  
RX MEDLINE=92052176; PubMed=194403;  
RA Slusher L.B., Gillman E.C., Martin N.C., Hopper A.K.;  
RT "mRNA leader length and initiation codon context determine  
RT alternative AUG selection for the yeast gene MOD5.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9789-9793(1991).  
RN [4]  
RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.  
RX MEDLINE=91203856; PubMed=1850093;  
RA Gillman E.C., Slusher L.B., Martin N.C., Hopper A.K.;  
RT "MOD5 translation initiation sites determine N6-Isopentenyladenosine  
RT modification of mitochondrial and cytoplasmic tRNA.";



RL Mol. Cell. Biol. 11:2382-2390(1991).

RN [5]

RP ALTERNATIVE INITIATION, AND SUBCELLULAR LOCATION.

RX MEDLINE-94187700; PubMed-8139355;

RA Boguta M., Hunter L.A., Shen W.C., Gillman E.C., Martin N.C.,

RA Hopper A.K.;

RT \*Subcellular locations of MOD5 proteins: mapping of sequences

RT sufficient for targeting to mitochondria and demonstration that

RT mitochondrial and nuclear isoforms commingle in the cytosol.;

RL Mol. Cell. Biol. 14:2298-2306(1994).

CC -1- FUNCTION: RESPONSIBLE FOR THE MODIFICATION OF A37 TO ISOPENTENYL

CC A37 OF BOTH CYTOSOLIC AND MITOCHONDRIAL TRNAs.

CC -1- CATALYTIC ACTIVITY: Isopeentenyl diphosphate + tRNA = diphosphate +

CC tRNA containing 6-isopentenyladenosine.

CC -1- PATHWAY: BIOSYNTHESIS OF THE MODIFIED BASE ISOPENTENYLADEOSINE

CC IN TRNAs.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL, CYTOPLASMIC AND NUCLEAR.

CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS MAY BE PRODUCED FROM THE USE OF

CC ALTERNATIVE INITIATION CODONS IN THE SAME READING FRAME.

CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: M15991; AAA34785.1;

DR EMBL: X89633; CAA61780.1;

DR EMBL: 275182; CAA99499.1;

DR PIR: A26711; A26717.

DR SGD: S0005800; MOD5.

DR InterPro: IPR002627; IPRF.

DR Pfam: PF01715; IPR1.

DR ProDom: PD004674; IPR1.

DR TIGRFAMs: TIGR00174; miaa; 1.

KW Transferase; tRNA processing; ATP-binding; Alternative initiation;

KW Mitochondrion; Nuclear protein.

FT CHAIN 1 428

FT TRNA ISOPENTENYLTRANSFERASE, CYTOPLASMIC/

FT MITOCHONDRIAL ISOFORM.

FT TRNA ISOPENTENYLTRANSFERASE, CYTOPLASMIC/

FT NUCLEAR ISOFORM.

FT FOR CYTOPLASMIC/NUCLEAR ISOFORM.

FT DOMAIN 210 232

FT BINDS ISOPENTENYLPYROPHOSPHATE

FT TRANSFERASE.

FT CONFLICT 313 313

FT MISSING (IN REF. 1).

FT CONFLICT 375 375

FT C -> R (IN REF. 1).

SQ SEQUENCE 428 AA; 50236 MW; A956B17ABC05161F CRC64;

Alignment Scores:

Pred. No.: 6,94e-35 Length: 428

Score: 548.00 Matches: 143

Percent Similarity: 52.58% Conservatve: 81

Best Local Similarity: 33.57% Mismatches: 144

Query Match: 15.33% Indels: 58

DB: 1 Gaps: 15

US-09-513-151-3 (1-2041) x MOD5\_YEAST (1-428)

OY 56 CTCAGGGGCGCTGCAACGGAGCCTACCTCTTGATGATCTCGGGCCGACGACGAC 112

DB 6 LeuLysGlyCysLeuAsnMetSerLysValIleValIleAlaGlyThrThrGlyVal 25

OY 113 GGCAAAATCCACGCTGCGCTGAGCTAGGCGCAGCGGCTCGGCGTGGAGTCTGACGCT 172

DB 26 GlyLysSerGlnLeuSerIleGlnLeuAlaGlnLysPheAsnGlyGluValIleAsnSer 45

OY 173 GACTCCATGACGCTATGACAGCGCTAGACATCATCCACACAGGTTTCTGCCAAGG 232

DB 46 AspSerMetGlnValLysLysAspIleProIleIleThrAsnLysHisProLeuGlnGlu 65

OY 233 CAGAGATCTGCCGGCAGCACATGATCAGCTTTGTGATCCTCTGTGACCAATTACACA 292

DB 66 ArgGlnGlyIleProHisIleValMetAsnHisValAsp---TrpSerGlnGlyTyr 84

OY 293 GAGGTGATCTCAGAAATAGAGCAACTGCTCGATGATGATATATATATATATATATAT 352

DB 85 SerHisArgPheGlnThrGlnCysMetAsnAlaIleGlnAspIleHisArgArgGlyLys 104

OY 353 ATTCCTATTGTTGTGGGAGAACCAATTATTCATTTGATCTGCTGTGCAAGTTCTT 412

DB 105 IleProIleValIleGlyIleGlyThrHisTyrTyrLeuGlnThrLeuPheAsnLys---Arg 123

OY 413 GTCAATCCACAGCCCGAGAGATGGC---ACTGAAAGTATATGACCGAAAGTGAG 469

DB 124 ValAspThrLysSerSerIleArgLysLeuThrArgLysGlnLeuAsp-----Ile 140

OY 470 CTGGAAGAGGAGATGGTTGTACTTACAAACGGCTTAAGCCAGGTGAGCCAGAAATG 529

DB 141 LeuGlnSerThrAspProAspValIleTyrAsnThrLeuValLysCysAspProAspIle 160

OY 530 GCTGCCAAGCTGCATCCACATGACAAACGCAAGTGCCAGAGCTTGCAAGTTTGA 589

DB 161 AlaThrLysTyrHisProAsnAspTyrArgValGlnArgMetLeuGlnIleTyrTyr 180

OY 590 GAAACAGGAATGCTCATGATGATTTCCATTCGTCACATACGGAAGAAGTGGTGT 649

DB 181 LysThrGlyLysLysProSerGlnThrPheAsnGlnLysIle----- 195

OY 650 CCCCTTGAGAGCTCTGAGTCTGACCTTACCTTCATCCCTTGATGCTGATGACGAC 709

DB 196 -----ThrLeuLysPhe---AspThrLeuPheLeuThrLeuTyrSerLysPro 210

OY 710 GCAGTTCTAGATGAGCGCTTGATAGAGGGTGGATGATGATGCTGCTGCGGCTTTG 769

DB 211 GluProLeuPheGlnArgLeuAspArgValAspAspMetLeuGlnArgGlyAlaLeu 230

OY 770 GAGGAACTAGAGATTTTACAGACGCTTATTCAGAAATGTTCCGAAATAGCCG 829

DB 231 GlnGlnIleLysGlnLeuTyrGlnTyrSerGlnAsnLysPheThr-----ProGlu 248

OY 830 GACTATCAACATGATCTTCCATCAATGATGCTTCAAGATTTTCACAGATGATGATC 889

DB 249 GlnCysGlnAsnGlyValIleProIleValIleGlyPheLysGlnPheLeuProIlePhe 267

OY 890 ACTGAGGGA---TGC----- 904

DB 268 ---ThrGlyLysThrAspAsnThrValLysLeuGlnAspCysIleGlnArgMetLys 286

OY 905 -----ACACTGGAGACTGATTAACCAAGCTTCRAAGAAAGACCTGGTCCATTGCCCC 958

DB 287 ThrArgThrArgGlnTyrAlaLysAspGlnValLysTyrPheLysLysMetLeuIlePro 306

OY 959 CCTGTCTATGCG-----TTAGAGTATGATGATGCTGAGAGGGAGAGATCT 1006

DB 307 AspIleLysGlyAspIleTyrLeuLeuAspAlaThrAspLeuSerClnThrAspHisn 326

OY 1007 GTTCTGAACCTGCTCTTGAATCGTCAAACTTCATC-----CAG 1048

DB 327 AlaSerGlnArgAlaIleAlaIleSerAsnAspPheIleSerAsnArgProIleLysGln 346

OY 1049 GGCCACAGCCCTACAGCCCACTCCATMAAGATGCCATACATGAGAGCTGAGAACAGAA 1108

DB 347 GluArgAlaProLysAlaLeuGlnGlnLeuLeuSerLysGlyLysThrThrMetLysLys 366

OY 1109 -----AGTTATCACCTGTGTGACCTCTGT-----GATGCA 1138

DB 367 LeuAspAspThrPheThrHisTyrThrCysAsnValCysArgAsnAlaAspGlyLysAsnVal 386

OY 1139 ATCATCATTTGGGATGGCAATGGCAGCAGCCACATMAATCCCACTTGAACAA 1198

DB 387 ValAlaIleGlyGlnLysTyrTyrLysIleHisLeuGlySerArgArgHisLysSerAsn 406

QY 1199 CTGAGAAAGAGAGA 1216  
 DB 407 LeuysArgAsnThrArg 412  
 RESULT 2  
 MIAA\_BACHD STANDARD: PRT: 314 AA.  
 AC Q9KAC3: 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPTP transferase) (IPTP)  
 DE (IPTase) (IPTP)  
 GN MIAA OR BH2366.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis."  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF 2-METHYLTHIO-N6-(DELT(2))-ISOPENTENYL-ADENOSINE (MS121(6)a))  
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.  
 CC -1- SIMILARITY: BELONGS TO THE IPTP TRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL: AP001515; BAB06085.1;  
 DR InterPro: IPR002627; IPTP.  
 DR Pfam: PF01715; IPTP: 1.  
 DR Prodom: PD004674; IPTP: 1.  
 DR TIGRfam: TIGR00174; miaa; 1.  
 KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding; Complete proteome.  
 FT NP\_BIND 10  
 FT SEQUENCE 314 AA; 35763 MW; E1027425C152942 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.01e-24 Length: 314  
 Score: 417.00 Matches: 102  
 Percent Similarity: 56.27% Conservative: 64  
 Best Local Similarity: 34.58% Mismatches: 85  
 Query Match: 11.66% Indels: 44  
 DB: 1 Gaps: 10  
 US-09-513-151-3 (1-2041) x MIAA\_BACHD (1-314)  
 QY 83 CTGTAGATGATCTCGGGCCACGGGACCGCAATTCACGCTGTCGACGTAGGC 142  
 DB 5 LeuValAlaIleValIglYrThrAlaValIglYsThylsTherValmetLeuAla 24  
 QY 143 CAGCGGCTCGCGGTGAGATGTCGACGCTGCTCATGAGAGCCCTAGAC 202  
 DB 25 LysArgLeuAsnIglValIleSerIglYAspSerMetGlnValIlyrArgIglYmetAsp 44  
 QY 203 ATCATCCACCAAGAGTTCTGCGCCCAAGACAGAGATCTGCGCGACCATGATCAGC 262

DB 45 IleglYThrAlaIleValIleThrAlaIglUglUmetAspIglYValProHisIleuIleAsp 64  
 QY 263 TTGTGATCCTCTTGTGACCAATTACACAGTGGTGGACTTCAGAAATAGACCACTGCT 322  
 DB 65 IleLysAspPro---SerGluSerPheSerValAlaAspPheGlnAspLeuAlaThrPro 83  
 QY 323 CTGATGAAATATATTGTCGCCAGACAAATTCCTATTGCTGTGGGGAGGAACCAATTAT 382  
 DB 84 LeuIleThrGlnIleHisGlnIglYrIglYrIglYrIglYrIglYrIglYrIglYrIglYr 103  
 QY 383 PACTATGATCTCTGCTGGAAGT---CTTGCATATACCAACCCAGAGATGGGC 439  
 DB 104 TyrValAsnAlaValIleHisGlnPheAsnLeuGlyAspIleArgAlaAspIleu----- 121  
 QY 440 ACTGAGAAAGTATGACCGAAGAGTGGAGCTTGA-----AAGAGAGATGCTT 490  
 DB 122 -----AspYrArgHisGlnIleuGlnAlaPheValAsnSerTyrIglYal 136  
 QY 491 ---GTACTTCACAAACGCTTAAGCCAGGTCGACCCAGAAATGCTGCCAAGCTGATCA 547  
 DB 137 GlnAlaLeuHisAspLysLeuSerLysIleAspProLysAlaAlaIleHisIsp 156  
 QY 548 CATGACAAAGCAAGAGTGGCCAGAGCTTGCAGAGTTTGAAGAAACAGAAATCTCTCAT 607  
 DB 157 AsnAsnTyrArgArgValIleArgAlaLeuGlnIleLysLeuThrIglYsThrVal 176  
 QY 608 AATGATTTCTCCATGTCATACATACGAAAGAGTGGTGGCTTGGAGCTCTG 667  
 DB 177 ThrGlu-----GlnAlaArgHisGlnIglYrIglYrIglYrIglYrIglYrIglYrIglYr 190  
 QY 666 AAGTCTCTACCTGATCCTTGGCTTCATGCTGACACGAGCATTCATAGAGCGC 727  
 DB 191 -----AsnLeuValMetIleGlyLeuThrMetGlnArgAspValLeuYrAspArg 207  
 QY 728 TGTGATAGAGGTGATGATACATGCTGCTGCTGGCTGTGGAGTGAAGTATGATATT 787  
 DB 208 IleAsnArgArgValAspIleMetValIglUglYleuIleAspIleuAlaLysIleu 227  
 QY 788 CACAGACGTATATACAGAAAGATGTCGAAATATAGCAGATATCAACATGCTATC 847  
 DB 228 -----TyrAspArgIlyle 232  
 QY 848 -----TTCCATCAATTGGCTTCAAGAAATTCACGAGTACTGATCACT 892  
 DB 233 ArgAspCysGlnSerValGlnAlaIleGlyTyrIglUmetYrAspYrIleu----- 250  
 QY 893 GAGGGAATATGACACTGAGACTGTATACAGCTTCTTAAGAA 937  
 DB 251 AspGlyAsnValThrLeuGlnIglUAlaIleAspThrLeuLysArg 265  
 RESULT 3  
 MIAA\_BACSU STANDARD: PRT: 314 AA.  
 AC Q31795: 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPTP transferase) (IPTase)  
 DE (IPTase) (IPTP)  
 GN MIAA.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertoletti M.G., Bessières P., Bolotin A., Borchert S., Bortiss R., Boursier L., Brans A., Braun M., Briganti S.C., Bron S., Brouillet S., Brunsch C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,

QY	260	AGCTTTTGATCTCTCTTGAGCAATTTACACAGTGGTGGACATTCGAAATTAAGCAACT	319
Db	67	AspIleuLeuAspPro---GlnAspSerPheSerThrAlaAspArgInSerLeuValArg	85
QY	320	GCTGTATGTAAGATATATTATGGCCGAGACAAATATCTATTGTGTGGAGGAAACCAAT	379
Db	86	AsnYsIleSerGluIleAlaAsnArgGlyYysLeuProMetIleAspGlyIleThrGly	105
QY	380	TATTACATTGAATCTGCTCTGGAAAGTCTTGTCATATACCAAGCCCGAGAGATGGGC	439
Db	106	LeuYrIleGlnSerGluLeuTyrAspTyrThrPheThr-----GluGluAla	121
QY	440	ACTGAGAAAGGATTTAGCCGAAAGCTGGAGCTT-----GAAAGAGGATGGCTTGTA	493
Db	122	AsnAspProValPheArgGluSerMetGlnMetAlaAlaGluArgGluGlyAlaAspPhe	144
QY	494	CTTCACAAAGCCCTTAAGCCAGGTGGACCCGAGAAATGGCTGGCCAGCTGCATCCACATGAC	553
Db	142	LeuHsiAlaIleAlaValAlaAlaAlaAspProGluAlaAlaAlaIleHsiProAsnAsn	161
QY	554	AAACGCCAAATGGCCAGAGCTTGCAGAGTTTTCGAAAGAACAGCAATCTCATATGTGA	613
Db	162	ThrTrgArgValIleAlaArgAlaLeuGluIleLeuHsiThrSerGlyLysThrMetSerGln	181
QY	614	TTTTCATCGCTCAACATACGAGGAAGAGTGTGTGCTCCCTGGAGAGTCCCTGTAAGTTC	673
Db	182	HsiLeuYsGlnGlnIlyAsnGln-----LeuLeu	191
QY	674	TCTAACCCCTTGACATCTTGGCTTCATGCTGACGAGCAGCTTCATGATGAGCCGCTGGAT	733
Db	192	TyrAsnAlaValLeuIleGlyLeuThrMetAspArgAspThrLeuTyrGluArgIleAsn	211
QY	734	AAGAGGTGATGATGATGATGCTGCGCGGCGTCTGGAGGAACGAAAGATTTTCAAGA	793
Db	212	GlnArgValAspLeuMetMetGlnSerGlyLeuLeuProGluVal-----Lys	222
QY	794	CGCATATATCAGAAATGTTTTCGGAANAATAGCCAGGACTATCAACATGTGATCTTCGAA	853
Db	228	ArgLeuTyrAspLysAsnVal-----ArgAspCysGlnSer-----IleGln	241
QY	854	TCAATTTGGCTTCAAGGAATTTTCACAGCATCTGATCACTGAGGAAATGCACACTGGAG	913
Db	242	AlaIleGlyTyrLysGluLeuTyrAlaTyrPhe-----AspGlyPheValThrLeuSer	259
QY	914	ACTAGTACACGAGCTTCAAGGAA	937
Db	260	AspAlaValGlnGlnLeuLysGln	267
RESULT 4			
MIAA_LISMO			
ID	MIAA_LISMO	STANDARD:	PRT: 305 AA.
AC	084713		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	tRNA delta(2)-Isopentenyl diphosphate transferase (EC 2.5.1.8) (IPP		
DE	transferase) (Isopentenyl diphosphate transferase)		
DE	(IPPTase) (IPPT)		
GN	MIAA OR LMO1294.		
OS	Listeria monocytogenes.		
OC	Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.		
OX	NCBI_TaxID=1639.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRATIN-BGD-e / Serovar 1/2a;		
RX	MEDLINE=21537279; PubMed=11679669;		
RA	Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,		
RA	Baquero F., Berend P., Bloeker H., Brandt P., Chakraborty T.,		
RA	Charbit A., Chetoui F., Couve E., de Daruvar A., Dehoux P.,		
RA	Domann K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,		
RA	Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,		
RA	Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,		

Alignment Scores:	
Pred. No.:	3,31e-23
Score:	397.50
Percent Similarity:	54.51%
Best Local Similarity:	34.03%
Query Match:	11.12%
DB:	1
	Gaps: 7
	Indels: 31
	Mismatches: 100
	Conservative: 59
	Matches: 305
	Length: 305

CC modified and this statement is not removed. Usage by and for commercial

entitles requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

-----

CC EMBL: AL596168; CAC96563.1; -

DR Listlist: L1N01332; -

DR InterPro: IPR002627; IPT.

DR Pfam: PF01715; IPT; 1.

DR ProDom: PD004674; IPT; 1.

DR TIGRFAMs: TIGR00174; miaa; 1.

KW Transferase: Nucleotidyltransferase; tRNA processing; ATP-binding;

KM Complete proteome.

FT NP\_BIND 11 18 ATP (POTENTIAL).

SO SEQUENCE 305 AA; 34691 MW; 1E46250F3BA78C07 CRC64;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4.73e-23	395.50	51.71%	33.22%	11.06%	305	97	54	102	39	6

US-09-513-151-3 (1-2041) x MIAA\_LISIN (1-305)

77 CTACCTCTGTAGTATCTCGGGCCAGCGGACCGGAAATCCACGCTGGGCTGACG 136

4 IIEPVALILEVALILEVALIGYPTOTHRALAVAGLYLSTHRSERIEUSERTLEHR 23

137 CTAGGCCAGCGGCTCGCGGTGATGATGTCAGCGCTGATGTCAGGCTATGAAGC 196

24 LEUVALALYASNPHEASPLGLUITLEISERGLYASPMETCLINVALTYRARGLY 43

197 CTAGACATATACCAACAAGCTTTTGCCCAAGACACAGAAATCTGCCGACCCATG 256

44 LEASPLILEGLYTHRALYSLIETHRPROGLUGLMEASPOLYILEYSHISTYRLEU 63

257 ATCAGCTTTGTCGATCTCTTGACCAATTAACACAGTGGTGGACGAATAAGACA 316

64 IIEASVALTHRASPROALAVAIPro---PHETHALALALYSPHEGLINALAGLUTHR 82

317 ACTGCTGATGATGAGATATATTGGCCGACGACAAATTCATTTGTTGGAGAGACC 376

83 ARGGLYLEUITLEGLUSERILEHNSAENARGGLYLEUENPROILELLEVALIGLYLTHR 102

377 AATTATTAATGATCTCTGCTGGAAGTTCTTGTAAATRCACAGCCCAAGAGATG 436

103 GLYLEUITYRILEGLINSEYVALPHEITYRASYGLYPHEGLYASNALASERGLUASPLYS 122

437 GGCAGTGAAGAAGTATGACCGAANAAGTGAGCTGAAGAAGAGATGGTCTTACTT 496

123 ALATYRATGATG-----GLULEASPLINLEUASPLYSTHRTHTLEU 136

497 CACAAAGCGCTAAGCCAGGTGACCCAGAAATGCTGCAAGTGCATCCATCAACA 556

137 TTPGLMETLEUASPLINLEUASPRPOLYSSERIALAGLULEUITLEHSLUNASNAHLS 156

557 CGCAAGTGGCCAGAGCTTGAAGTTTGAAGAAGAACTCTCTATAGTGAATTT 616

157 ARGARGVALIILEHGLALEUGLIVALLIENHLSLEUTHGLIYASPROPHESERGIUUY 176

617 CTCGACATGCAATACGAGAGAGAGTGGTGGTGGAGTGGAGTCTGTAAGTTCTGT 676

177 GLINVALIHISHTHLEUASNGULALATYRGLNPROLEU----- 189

677 AACCCCTTGATCTCTTGCTCATGCTGACGAGGAGCTTGTAGATGAGCGCTGATAG 736

190 -----PHELEUGLYLEUASPLINLEUASPRARGLULEUENLEUTHYRGLINARG 206

737 AGGGTGAATGACATCTCTGCTGCTGCTTGGAGAACTAAGAGATTTTCACAGAGCC 796

207 ARGVALIGLULEUENPHEGLUGLUGLYLEU----- 216

797 TATAATCAGAGAAATTTTCGAAATATAGCAGGACTATCAACATGTATC----- 847

-----

DB 217 -----VALSERGLUASERLYSLEUTHYRASPGLUINHSLEUVALSPAL 231

848 -----TTCCATCAATTTGGCTTCAAGAAATTCACAGTACTGATCAAGGAGAA 901

232 PROALILEARGLYLILEGLYRYSGLUENPHEHTYRPH-----ASPGLYSN 249

902 TGCACACGTGACATGATACAGGCTCTCAAGAA 937

250 SERSERLEUGLUGLUALYSGLUENLEUGLINS 261

RESULT 6

MIAA\_CLOAB STANDARD: PRT; 309 AA.

ID MIAA\_CLOAB

AC 097121:

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase)

DE (IPPT)

GN MIAA OR CAC1835.

OC Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

OX NCBI\_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J., Qiu D., Hitt J., Wolf Y.I., Tatusov R.V., Sabathe E.V., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;

RA "Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum."

RT J. Bacteriol. 183:4823-4838(2001).

RL - FUNCTION: Catalyzes the first step in the biosynthesis of 2-methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS[2]I[6]A) adjacent to the anticodon of several tRNA species (By similarity).

CC - CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate + tRNA containing 6-isopentenyladenosine.

CC - SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).

CC -----

CC EMBL: AE007692; AAK79799.1; -

DR InterPro: IPR002627; IPT.

DR Pfam: PF01715; IPT; 1.

DR ProDom: PD004674; IPT; 1.

DR TIGRFAMs: TIGR00174; miaa; 1.

KW Transferase: Nucleotidyltransferase; tRNA processing; ATP-binding;

KM Complete proteome.

FT NP\_BIND 9 16 ATP (POTENTIAL).

SO SEQUENCE 309 AA; 35687 MW; 9330A81B46E8AC3E CRC64;

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1e-21	378.50	54.64%	34.02%	10.59%	309	99	60	99	33	9

US-09-513-151-3 (1-2041) x MIAA\_CLOAB (1-309)

83 CTGTAGTATCTCTCGGGCCAGCGGACCGGAAATCCACGCTGGCTTGCAGCTAGGC 142



[illegible]

CC		EMBL; AL64670; CAD16271.1;	-
DR		InterPro; IPR002627; IPT.	
Dr	Pfam:	PF01715; IPT; 1.	
DR	ProDom:	PD004674; IPT; 1.	
DR	TIGRFAMS:	TIGR00174; miaA; 1.	
KW	Transferase,	Nucleotidyltransferase; tRNA processing; ATP-binding;	
KM	Complete proteome.		
FT	NP_BIND	16                  23                  ATP (POTENTIAL).	
SQ	SEQUENCE	323 AA; 34966 MW; DP0086716A54E180 CRC64;	
Alignment Scores:			
	Pred. No.:	1.63e-20	Length: 323
	Score:	363.00	Matches: 105
	Percent Similarity:	49.68%	Conservative: 48
	Best Local Similarity:	34.09%	Mismatches: 114
	Query Match:	10.15%	Indels: 41
DB:		Gaps: 1	9
US-09-513-151-3 (-1-2041) x MIA_RALSO (-1-323)			
OY	64	CCTGCAGACGACCCTTACTGTATGATTTCGGCGCACGCCGACAACCGCAATCCAC	122
Dd	5	PROHSHIHAIAPROARAla ValCysleuLeuLyrproThralaserelystrhal	24
OY	124	GCTGCGTTTGACGTAGGCCACGCGGCTGGGGGTGAGATGTCACGCCCTCATTCATGA	183
Dd	24	aalalaaleualaaleualaglnargtrPrrovalglunlleisermetaseralale	44
OY	184	GGTATATGAGGCCCATGACATCATCCACAAACAGTTCTGCCAACAGCAGAGAATCTG	243
Dd	44	vallaylrarpsmetsapilegtylthlalalytprosergammaaglucinalatalal	64
OY	244	CCGACACACATGATCAGACTTGTGGATCCTCTTGACCAATTAACAGAGTGGACTT	303
Dd	64	aproHISHIleuulleaspilleleasprroleu---Aspaleryserlaalaglnph	83
OY	304	CAGAAATAGACCACTGCTCTGATTTGAAGATATTTGGCCGAGACAAAATTCCTATGT	363
Dd	83	ealathraspaIngalnataleuulleagualtalarygaalarglyrgleuPreoleuil	103
OY	364	TGTGGAGAGAACCAATTATTAACATTGATCTGCTCGGAGAAAGTTCCTGCAATACAA	423
Dd	103	evalcllyglyhmletuleuyr-----Tyrlsalaleuthrhnglnyle	118
OY	424	GCCCCAGAGATGGGCACTGAGAAAGTAGATTGACCGAAMAATGGAGCTTGAA--AAGA	480
Dd	118	userserpleurProgilaalsaprohalalleargAlaleagluilleasppladguallalear	138
OY	481	GGATGGTCTT---GIACITTCACAAAGCCTTAGACCCAGGTGAGCCCAATAATGCTGCCA	537
Dd	138	gaspeLyrrPrroalaleuuhlsatalsvsleuaialglvlasprrovalthrllaalaar	158
OY	538	GCTGCATCCACATGACCAACCAAGTGCGCAGAGAGCTGCAOATTTTTGAAGAACAG	597
Dd	158	gleuhtslalahrsapaIngarlgllleglnargAlaleuclsleuueyryrgdleuthrei	178
OY	598	AATCTTCATAGTGATTTCTCCATCGTCAA-----CATACGGAGA	639
Dd	178	yglNprometserialeuuleualatargduallaiglAlalaalarphenhsilarginisgi	198
OY	640	AGGTGGTGGTCCCTGGAGAGCTCTGGAAGTTCACACCCCTTGACCTTTGGCTCA	699
Dd	198	uAlaAlaAla-----AlrtyleusertlleAleugltupr	210
OY	700	TGCTGACACGACCTTAGATGAGCGCTTGATTAAGAGGAGGTGCATACATCTTGCTGC	759
Dd	210	oAlasprgammaAlavalleuhtsisaatrgllleaglnArgphensprlametleualagi	230
OY	760	TGGCTCTTGAGAGAACTAAAGATTTTCAACAGACSTATAATCAGAAAGATTTTCGGA	819
Dd	230	yellyleuueuspcluvallAlualeuarayagaaggilyAspleuserProvalleupr	250



QY 820 AATAGCCAGACTATCAACATGTATCTTCATTCAGTTCGCTTCAAGAAATTCACGA 879  
 Db 250 oser-----1leargcysvalglytyrarglnala1trpal 262  
 QY 880 GTACTGATCACTAGAGGAAATGACACTGAGACT----- 916  
 Db 262 atyrleu-----aspglcyluileaspmetalatrrleuarglucinglylleala1 280  
 QY 917 -AGTAACGAGCTCTTAAGAAA 937  
 Db 280 arthrarglnleucyslysarg 287  
 RESULT 9  
 MIAA\_PSEPU STANDARD: PRT; 322 AA.  
 AC O30762;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP  
 DE transferase) (isopentenyl-diphosphate:trna isopentenyltransferase)  
 DE (IPPT)  
 DE MIAA.  
 OS Pseudomonas putida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M;  
 RA Olekhovich I.N., Gussin G.N.:  
 RT "Attenuation of the Pseudomonas putida type and tirpdc genes."; Submitted (Jul-1997) to the EMBL/GenBank/DBJ databases.  
 RL -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF  
 CC 2-METHYLLITHIO-N6-DELTA(2)-ISOPENTENYL-ADENOSINE (MS1211[6]A)  
 CC ADACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: isopentenyl diphosphate + trna -> diphosphate +  
 CC trna containing 6-isopentenyladenosine.  
 CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF016312; AAB69443.1; -  
 CC InterPro: IPR002627; IPPT.  
 CC Pfam: PF01715; IPPT; 1.  
 CC ProDom: PD004674; IPPT; 1.  
 CC TrEMBL: TIGR00174; miaa; 1.  
 CC Transferase; Nucleotidyltransferase; trna processing; ATP-binding.  
 FT NP\_BIND 12 19 ATP (POTENTIAL)  
 FT SEQUENCE 322 AA; 35488 MW; DIAA81AD7B32F6A CRC64;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2e-19 Length: 322  
 Score: 349.00 Matches: 89  
 Percent Similarity: 53.248 Conservative: 67  
 Best Local Similarity: 30.388 Mismatches: 109  
 Query Match: 9.768 Indels: 28  
 DB: 1 Gaps: 9  
 US-09-513-151-3 (1-2041) x MIAA\_PSEPU (1-322)  
 QY 80 CCTCTGTAGATGATTCCTCGGGCCAGCGGCAACGCAATTCACGCTGGCGTTGCAGCTA 139  
 Db 6 PROAlaIlePheLeuMetGlyProThrAlaIleAglGlyThrAspLeuAlaIleGluLeu 25  
 QY 140 GGCACAGGCTCGCGCTGAGATCGTCACGCGCTGATCCATCGACGCTCTGAGGCTTA 199  
 Db 140 GGCACAGGCTCGCGCTGAGATCGTCACGCGCTGATCCATCGACGCTCTGAGGCTTA 199

Db 26 ThrIysValLeuProCysGluLeuIleSerValAspSerAlaLeuValTyrIrrgIlyMet 45  
 QY 200 GACATCATCACAACAAGGTTTCTGCCAAGACAGAGAATTCGCCGACCATGATC 259  
 Db 46 AspIleGlySerAlaLysProSerLysGluIleLeuAlaAlaHisProHisArgIleLeu 65  
 QY 260 AGCTTTGGTGGATCCTTCTTGACCAATTCACAGCTGCTGAGCTGCAATTCAGCAACT 319  
 Db 66 AspIleArgAspPro--AlaGluSerTyrSerAlaAlaGlnPheArgAlaAspAlaLeu 84  
 QY 320 GCTCTGATGAAAGATATATTTGCGCGAGCAAAATTCATTTGTTGGAGAGCAACT 379  
 Db 85 GluAlaMetAlaGluIleThrAlaArgGlyLysIleProLeuLeuValGlyIleThrMet 104  
 QY 380 TATTACATGGAATCTCTGCTCTGGAAAGTTCTTGTCAATACCAAGCCCAAGAGATGGC 439  
 Db 105 LeuTyrTyrLysAlaLeuIleAspGlyLeu-----AlaAspMetPro 118  
 QY 440 ACTGAGAAAGTGATGACCGGAAAGTGAGACTGAAAGAGGAT--GGTCTGTA--- 493  
 Db 119 AlaAspAlaAlaValAlaArgAlaGluLeuGluArgGlnAlaGluAlaLeuGlyLeuAlaGlu 138  
 QY 494 CTTCACAAAGCGCTTAAGCCAGGTGACCCAGAAATGCTGCCAAGCTGCATCCATGAC 553  
 Db 139 LeuHisArgGlnLeuAlaGluValAspProGluSerAlaAlaArgIleHisProAsnAsp 158  
 QY 554 AAAGCAAAAGTGCCAGGAGCTTGCAGATTGTTGAAGAAACAGAAATCTCTACTGAA 613  
 Db 159 ProGlnArgLeuIleArgAlaLeuGluValTyrArgValSerGlyLysMetThrAla 178  
 QY 614 TTTCCTCCATCGTCAACATACGGAAGAGGTGGGTCC-----CTGGAGGCTCTG 667  
 Db 179 HisArgGlnArgIlePheAlaGluSerArgGlyAlaAspAlaGlyAlaGlyLysIleLeu 198  
 QY 668 AAGTCTCTAACCTTGATCCTTGGCTTCAAGCTGACCGACGACGACTTGTAGTAGCC 727  
 Db 199 ProTyrThrValAlaSerLeuAlaIleAlaProThrAspArgHisIleLeuHisGlnArg 218  
 QY 728 TTGGATPAGAGGTGATGATGATGATGCTGCTGCTGGCTCTTGAGAGAACTAGACTTTT 787  
 Db 219 IleAlaLeuArgPheSerGlnMetLeuGluGlnGlyPheValAspGluValArgSerLeu 238  
 QY 788 CACAGACGCTATATATCAGAAAGATGTTTGGAAATACGACATCATCAACTGATC 847  
 Db 239 ArgAlaArg-----SerAspLeuHisAlaGlyLeu 248  
 QY 848 -----TTCCATCAATTTGGCTTCAAGAAATTCACAGATCTGATCTGAGGAGAAA 901  
 Db 249 ProSerIleArgAlaValGlyTyrArgGlnValTyrAspTyrLeu-----AspGlyLys 266  
 QY 902 TGCACACTGAGACTAGTAACGACTCTTAAGAAAAGA 940  
 Db 267 Leu-----ThrGluAsnGluMetArgGluArgGly 276  
 RESULT 10  
 MIAA\_THEME STANDARD: PRT; 305 AA.  
 AC O9WY25;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP  
 DE transferase) (isopentenyl-diphosphate:trna isopentenyltransferase)  
 DE (IPPT)  
 DE MIAA OR TM0525.  
 GN Thermotoga maritima.  
 OS Bacteria; Thermotogae; Thermotogae (class); Thermotogales;  
 OC Bacteria; Thermotogae; Thermotogae.  
 OX NCBI\_TaxID=2336;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MSB / DSM 3109;  
 RX MEDLINE=99287316; PubMed=10360571;



RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 RA McDonald L., Utterback T.R., Malek J.A., Lither K.D., Garrett M.M.,  
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.,  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 RT genome sequence of *Thermotoga maritima*,"  
 RL Nature 399:323-329(1999).  
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF 2-  
 CC METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]I[6]A))  
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA - diphosphate +  
 CC tRNA containing 6-isopentenyladenosine.  
 CC -1- SIMILARITY: BELONGS TO THE ITP TRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001728; AAD35610.1; -  
 DR TIGR: TM0525; -  
 DR InterPro: IPR002627; IPTP.  
 DR Pfam: PF01715; IPTP.1.  
 DR ProDom: PD004674; IPTP.1.  
 DR TIGRFAMS: TIGR00174; miaA; 1.  
 KM Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;  
 KM Complete proteome.  
 FT NP\_BIND 8 15 ATP (POTENTIAL).  
 SQ SEQUENCE 305 AA: 35926 MW: 5930845AEC64AC3 CRC64:  
 Alignment Scores:  
 Pred. No.: 2,366-19 Length: 305  
 Score: 348.00 Matches: 86  
 Percent Similarity: 53.66% Conservative: 68  
 Best Local Similarity: 29.97% Mismatches: 105  
 Query Match: 9.73% Indels: 28  
 DB: 1 Gaps: 8  
 US-09-513-151-3 (1-2041) x MIAA\_THEMEA (1-305)  
 QY 77 CTACGCTGTGATGATTCGCGGCCACGGCCGGAATCCACGCTGCGGTTGCAG 136  
 DB 1 MetLysIleAlaIleValaIleGlyProThrAlaValaIleGlyThrAspIleMetIleGlu 20  
 QY 137 CTAGGCCAGCGCGCTGGCGGTGAGATCGTCAGCGCTGACATGACGATGATGAGGC 196  
 DB 21 ValGysGluGluIleGlyAlaGluIleIleSerMetAspSerArgGlnIleIleTyr 40  
 QY 137 CTAGCATATCATCCACACAGATTTGCCCAAGACAGAGAAATGCGCGCACCATG 256  
 DB 41 MetAspIleGlyThrAlaLysProThrProGluGlnArgLysArgValIleHisMet 60  
 QY 257 ATCAGCTTTGGATGATCTTGTGACCAATTCACAGAGTGTCGATTCGAAATAGAGCA 316  
 DB 61 IleAspIleIleAspPro---AspGluTyrTyrAsnAlaPheMetTyrArgLysAspSer 79  
 QY 317 ACTGCTGTGATGAGATATATTGCGCGAGACAATTCATGTGTGGGAGAGACC 376  
 DB 80 LeuArgAlaMetGluAspValLeuArgArgLysIleProValTyrValGlyGlyThr 99  
 QY 377 AATTATTTCATGATTCCTCTGCTGGAAGTCTTGTCAATACCAAGCCCGAGAGATG 436  
 DB 100 GlyLeuTyrAlaAspAlaLeuVal---ArgGlyIlePheGluGlyValProAlaAspGlu 118  
 QY 437 GGCACGTAGAAAGTATGACCGAAATGACGCTGAAAAGAGAGATGCTTACTT 496  
 DB 119 AsnIleArgLysGluLeuArg-----GluLeuGlnArgArgGluProGlyIleLeu 135

QY 497 CACAAAGCGCTTAAGCCAGGTGACCCAGAAATGCTGCCAGACTGCATCCATGACAA 556  
 DB 136 ArgLysMetLeuGluIleuAspProGluAlaAlaThrArgIleHisProAsnAspLeu 155  
 QY 557 CCCAAAGTGGCCAGAGCTTGCAGATTTTGAAGAAACAGAAATCTCATGATGATTT 616  
 DB 156 LysArgThrIleArgAlaLeuGluValTyrMetLysThrGlyArgArgIleSerGluLeu 175  
 QY 617 CTCATCTGCAACATACGGAAGAGGTGGTCCCTTGAGAGTCTCTGATGTTCTCT 676  
 DB 176 -----GlnLysGluAlaLysGlyAspAsp-----ArgPhe----- 185  
 QY 677 AACCTTGATCTGCTTGGCTTCATGCTGACACGCGCTCTAGATGACGCTTGATAG 736  
 DB 186 -----PheIleIleValIleuThrArgGlyArgLysGluArgLysGluArgLys 203  
 QY 737 AGGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 796  
 DB 204 ArgValAspLysMetIleGluMetGlyLeuValAspGluValLysArgLeuGlyMet 223  
 QY 797 TTTAATCAGACAAATGTTTGGAAAAATACCGAGACTATCAACATGATTCCTCAATCA 856  
 DB 224 GlyTyrSerLysAspLeuAsnSer-----MetLysThr 234  
 QY 857 ATTGCTTCAAGCAATTCACAGATACGATGACGATGAGGAATGCACACTGAGACT 916  
 DB 235 ILeGlyTyrLysGluValIleAspTyrIleu-----GluGlyLysTyrAspAspLys 252  
 QY 917 AGTAACAGCTTCTTAAGAA 937  
 DB 253 MetValHisLeuIleLysArg 259  
 RESULT 11  
 MIAA\_CHLTR STANDARD: PRT: 314 AA.  
 AC 084771;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPT  
 DE transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase)  
 DE (Iprase) (IPTP).  
 GN MIAA OR CT766.  
 OS Chlamydia trachomatis  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UM-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;  
 RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis,"  
 RL Science 282:754-759(1998).  
 CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF  
 CC 2-METHYLTHIO-N6-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS[2]I[6]A))  
 CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA - diphosphate +  
 CC tRNA containing 6-isopentenyladenosine.  
 CC -1- SIMILARITY: BELONGS TO THE ITP TRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE001349; AAC68361.1; ALT\_INIT.  
 DR InterPro: IPR002627; IPTP.



```

OY 184 GGTCTATGAAGGCTAGCATCATCATCAACAAGTTTCTGCCCAAGACAGCAATCTG 243
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 41 uileiyrhaggllymetaspllelythrhlalsprosealaglgluldeuallaleuall 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 244 CCGGACACACATCAGCTTTGTGTGATCTCTTGTACCAATTTACAGTGTGACTT 303
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 aProHlSArGLlellAspRllecYasPro---AlaGlUsErTySeRlAlaLAsnRh 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 304 CAGAAATFACAGCACTGCTCTGATTGAGATATATTGCGCCGACAAATTCCTATTGT 363
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 80 eArvgInAspAlaLeuArGLlUmetAlaAspRllecAlaLaglLySllePrOleuLe 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 364 TGTGTGAGGAGAACATTTATTACATTCATCTGCTCTGGAAGTCTTGCTCAATCCAA 423
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 100 uValGlYglYThmleUerTyTyRlYsAlaLeu-----LeuGlYlYleuSe 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 424 GCCCCAGAGAGTGGCAGTGAAGAGTGTG-----ATTGACGGAAGAGTGGAGTGA 474
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 116 rProLeuPrOseRAlaAspGlUlYsValArGserGlUlleGlUlYlYsAlaLeu--- 135
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 475 AAAGGAGAGTGTCTGTCTACTTCACAAAGCCCTAACCCAGGTGAGCCCAAGATGCTGC 534
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 136 ---GInglYTPAlaAlaLeuHlSglInglUleuAlaLyslLeAspPrOleuAlaLagl 154
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 535 CAGAGTCATCCACATGACAAAGCAAGTGGCCAGAGCTTGCAAGTTTGAAGAAGC 594
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 154 nArGlEaSnPrOAsnAspSerGlInArGylEaSnHArGAlaLeuGlUlYAlaPhenYrLeuH 174
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 595 AGGAATCTCTCATAGTGAATTTCTCCATGCTCAACATACGGAAGAGTGTGTCTCCCT 654
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 174 rGlYlYsSerlEaSerGlUlEaSerGlInglInLys----- 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 655 TGGAGTCTCTGAGAGTCTCTCAACCTTGATCCCTTGCTGCTTCAAGTGCAGCAGCAGT 714
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 186 -cLyspSerlEaPrOtyrGlInlLeuGlInPhEaAlaLeaPrOlyAspArGserlI 205
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 715 TCTAGATGAGCCCTTGATGAAGAGGCTGATGATGCTTGCTGCTGCTGCTGAGGA 774
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 205 eLeuHlSAspArGlEaLeaLeuArpHeGlInLysHetiGleGlInglYPhEaGlIngl 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 775 ACTAAGAGATTTTTCACAGACGCTATATATGAGAAGATTTTGGAAAAATAGCCAGCACTA 834
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 225 uValGlUlYlYsLeuYrGlInArG-----GlUsAspLe 235
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 835 TCAACATGCTATC-----TTCCATCATTTGCTTCAAGAAATTTACAGACTACCTG-- 886
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 235 uHlSleuAspLeuPrOAlaMeLArGysValGlYrArGInMeTlPrGlUlYrLeuAr 255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 887 -----ATCACGAGGGAATGCACACTGAGACTAGTAA 921
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 255 gGlYAspTyRAspHlSAspGlUmetlEaRglYllecYs-----AlaThrAr 272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 922 CCAGCTTCTTAAGAAA 937
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 272 gGlInleuAlaLysArG 277
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

## RESULT 13

MIAA\_STAMM STANDARD: PRF: 311 AA.

```

AC 0990H0;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPP
DE transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase)
DE (IPPTase) (IPPT).
GN MIAA OR SAV1304 OR SALL144 OR MM1186.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620.

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Mu50 / ATCC 700699, and N315;
RX MEDLINE-21311952; PubMed-11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ul Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Mu2;
RX MEDLINE-22040717; PubMed-12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -1- FUNCTION: Catalyzes the first step in the biosynthesis of 2-
CC methylthio-N6-(delta(2)-isopentenyl)-adenosine (MS21[16A])
CC adjacent to the anticodon of several tRNA species (By similarity).
CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + tRNA -> diphosphate +
CC tRNA containing 6-isopentenyladenosine.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL; AP003361; BAB57466.1; -
CC DR EMBL; AP003133; BAB42398.1; -
CC DR EMBL; AP004826; BAB95051.1; -
CC DR InterPro: IPR002627; IPPT.
CC DR Pfam: PF01715; IPPT; 1.
CC DR ProDom: PD004674; IPPT; 1.
CC DR TIGRFAMs: TIGR00174; miaa; 1.
CC DR Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;
CC KW Complete proteome.
CC FT NP_BIND 13 ATP (POTENTIAL).
CC SO SEQUENCE 311 AA; 35868 MW; 6011E9D44878043E CRC64;

```

## Alignment Scores:

Pred. No.:	1.19e-18	Length:	311
Score:	339.00	Matches:	86
Percent Similarity:	55.64%	Conservative:	62
Best Local Similarity:	32.33%	Mismatches:	92
Query Match:	9.48%	Indels:	26
DB:	1	Gaps:	7

US-09-513-151-3 (1-2041) x MIAA\_STAMM (1-311)

```

OY 83 CTTGATGATGATTCGGGGCCACGGGACGCAATTCACGCTGCGTGGAGCTAGGC 142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 8 lIeValAllelValGlYlProHrhlAspRllecYlYsThrGlUleuSerllecGlUleuAla 27
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 143 CAGCGGCTCGGGGTGAGATGCTGACGCTGCTCATGCTGAGCTATGAGAGCCTAGAC 202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 28 lYsArGylEaSnGlYglInlEaSerGlYAspSerMetGlnValTyRlYsHlSmetAsn 47
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 203 ATCATCACCAACAGATGCTTCTGCCAAGAGCAGAGATCTGCCGGACCAACATGTATGAGC 262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 48 lIeGlYThrAlaLysValAlThrPrOglUglUmetAspGlYlIePrOHLlSleuIlEaSp 67
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 263 TTGTGATGCTCTTGTGACCAATTAACAGATGCTGACCAATTAACAGCACTGCT 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11eLeuAspProAspAspThr---PheSerAlaTyrGluPheLysArgLeuAlaGluAsp 86
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 323 CAGATGGAAGATATATGTCGCCGAGACAAATTCATGCTGTGCGAGGACCAATTAAT 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 LeuIleThrAspIleThrAsnArgGlyLysValProIleLeuAlaGlyIleThrGlyLeu 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 383 TACATGAATCTCTGCTGCGAAGATCTTGTCAATATACCAAGCCGAGAGATGGGCACT 442
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 TyrIleGlnSerIleuLeuIleTyrGlnLeuGlnAspIleValThrProAlaGln 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 443 GAGAAAGATATGACCGAAGAGTGGAG---CTTGAAGAGAGAGATGCTTGTACTTTCAC 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 127 LeuSerIleValLysGlnLysLeuSerAlaLeuGlnIleHisLeuAspGlnGlnLeuHis 146
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 500 AAGGCTTAAGCCAGATGAGCCAGAAATGCTGCCAAGCTGATCCACATGACAAAGCC 559
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 147 AspTyrIleAlaGlnPheAspAlaValSerAlaGlnAsnIleHisProAsnAsnArgGln 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 560 AAGTGGCCAGAGATGTCAGATTTTGAAGAAACAGAG---ATCTCTCATAGTGA 613
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 167 ArgValLeuArgAlaIleGluTyrIleuLysThrLysLysLeuLeuSerAsnArgLys 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 614 TTTCTCATGCTGACATACGAGAAAGAGTGTGCTCCCTTGAGAGCTCTGAACTTC 673
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 LysVal---GlnGlnPheThrGluAsn----- 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 674 TCATACCCCTTGATCTCTGCTGCTCATGCTGACAGGCACTTGTATGAGCGCTTGAT 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 195 TyrAspThrLeuLeuGlnGlyIleGluMetSerArgLysThrLeuTyrSerAlaGln 214
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 734 AAGAGGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 793
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 215 LysArgValAspIleMetLeuAspHisGlyLeuPheArgGlyVal----- 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 794 CGCTAAATCAGAGAAATGTTTCGAAATATGCCAGACATATACATGATCTTCCAA 853
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 230 -----GlnGlnLeuValGlnGlnGlyTyrGlnSerGlnSer-----MetGln 244
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 854 TCATATGCTTCAAGGA 871
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 245 AlaIleGlyTyrLysGln 250
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RT Rd":
RL Science 269:496-512(1995).
CC -1- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF
CC 2-METHYLNIO-M6-(DELTA(2))-ISOPENTENYL-ADENOSINE (MS12)1(6)A)
CC ADJACENT TO THE ANTICODON OF SEVERAL TRNA SPECIES (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate + trna -> diphosphate +
CC trna containing 6-isopentenyladenosine.
CC -1- SIMILARITY: BELONGS TO THE IPP TRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U32692; AAC21746.1;
CC TIGR: H10068;
CC InterPro: IPR002627; IPRP.
CC DR Pfam: PF01715; IPRP.
CC DR PRODOM: PD004674; IPRP.
CC TIGRFRAMS: TIGR00174; mlaB: 1.
CC KW Transferase; Nucleotidyltransferase; trna processing; ATP-binding;
CC FT complete proteome.
CC NP_BIND 10 17 ATP (POTENTIAL).
CC SQ SEQUENCE 311 AA; 35204 MW; 4B5E3EEFE28DC37 CRC64;

Alignment Scores:
Pred. No.: 2 67e-18 Length: 311
Score: 334.50 Matches: 93
Percent Similarity: 51.52% Conservative: 60
Best Local Similarity: 31.31% Mismatches: 103
Query Match: 9.36% Indels: 41
DB: Gaps: 11

US-09-513-151-3 (1-2041) x MIAA_HAEIN (1-311)
QY 86 GTAGTATCTCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 145
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 6 IlePheLeuMetGlyProThrAlaSerGlyLysThrAspLeuAlaIleGlnLeuArgSer 25
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 CGGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 26 GlnLeuProValGlnValIleSerValAspSerAlaLeuIleTyrLysGlyMetAspIle 45
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 206 ATCCACCAACAGGTTTCTGCCCAAGACAGAGAAATCTGCCGCGCACATGATCAGCTTT 265
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 46 GlyThrAlaLysProSerLysGlnGlnLeuAlaLeuAlaProHisArgIleAspIle 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 266 GTGGATCTCTGTGTGACCAATTAACAGTGTGAGTTCAGAAATAGAGCACTGCTCTG 325
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 66 LeuAspPro---SerGlnSerTyrSerAlaMetAsnPheArgAspAlaLeuArgGln 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 326 ATTCACATATATATTTGGCCCGAGACAAATTCCTATTGTTGGGAGAGAACCAATTATAC 385
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 85 MetAlaAspIleThrAlaGlnGlyLysIleProLeuLeuValGlyGlyThrMetLeuTyr 104
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 ATTCATCTCTGCTGCGAAGTCTGTGCAATATCAAGCCCGAGAGATGGGCACTGAG 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 105 TyrIleAlaLeu-----IleGlnLysLeuSerProLeuProSerAlaAspGln 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 446 AAGATGATGACCGAAGAGTGGAGCTTGA-----AAGAGAGATGCTCTTGTGA 493
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 AsnIle-----ArgAlaGlnLeuGlnGlnLysAlaIleAlaGlnGlnGlyTyrAlaAla 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 494 CTTGACAAAGCCCTTACCCAGAGTGGAGCCAGAAATGCTGCCAAGCTGCATCCAGACAGAC 553
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 138 LeuHisThrGlnLeuAlaLysIleAspProIleSerAlaAlaArgIleAsnProSerAsp 157
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 554 AAGCAGAAAGTGGCGGAGAGCTTCAAGTCTTGAAGAAACAGAGATCTCTCATAGTGA 613
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 158 SerGlnArgIleAsnArgAlaLeuGlnValPheTyrIleThrGlyLysSerLeuThrGln 177
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Whole-genome random sequencing and assembly of Haemophilus influenzae

QY 614 TTCTCCATGTCACATAGAGAGGNGGTGCTCCCTGGAGGTCTCTGAAGTTC 673  
 DB 178 Leu-----ThrluglnlysglygluAlaLeu-----ProTyrAspPhe 190  
 QY 674 TCTAACCTTGATCTTGGCTTCATGCTGACACGAGCATTCATAGAGCGCTTGAT 733  
 DB 191 ValGlnPheAlaIle-----AlaProGlnAspArgHisValLeuHisGlnArgIleGln 208  
 QY 734 AAGAGGCTGATGACATGCTGCTGCTGCTGCTGGAGAGACTAAGATTTTCACAGA 793  
 DB 209 GlnArgPheHisLysMetIleGlnLeuGlnAlaGlnValGlnLysLeuTyrAla 228  
 QY 794 CGCTATATACAGAGAGATTTGCGAAATTAAGCCAGACTATCAACATGATCTCCAA 853  
 DB 229 Arg-----GlyAspLeuAsnIleAsnLeuProSer-----IleArg 240  
 QY 854 TCATATGCTTCAGAGAAATTTTCACGAGTACTG----- 886  
 DB 241 CysValGlyTyrArgGlnMetTrpGluTyrLeuGlnGlyAspTyrAlaTyrGlnGlnMet 260  
 QY 887 ATCACTAGAGGAAATGCACACTGAGACTAGTACACCTCTTAAGAAA 937  
 DB 261 IlePheArgGlyIleCys-----AlaThrArgGlnLeuAlaLysArg 274

RESULT 15  
 M1AA\_VIBCH STANDARD: PRT: 315 AA.  
 AC 09KV12;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE tRNA delta(2)-isopentenylpyrophosphate transferase (EC 2.5.1.8) (IPT  
 DE transferase) (isopentenyl-diphosphate:tRNA isopentenyltransferase).  
 DE (IPIase) (IPIPT).  
 GN M1AA OR VC0346.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.V., Haft D.H., Hickey E.K., Peterson J.D., Unsay M.L.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -I- FUNCTION: CATALYZES THE FIRST STEP IN THE BIOSYNTHESIS OF 2-  
 CC METHYLNHE-NE-(DELTA(2)-ISOPENTENYL)-ADENOSINE (MS1211(6IA))  
 CC ADJACENT TO THE ANTICODON OF SEVERAL tRNA SPECIES (BI SIMILARITY).  
 CC -I- CATALYTIC ACTIVITY: isopentenyl diphosphate + tRNA = diphosphate +  
 CC tRNA containing 6-isopentenyladenosine.  
 CC -I- SIMILARITY: BELONGS TO THE ITP TRANSFERASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: AE004123; AAF93519.1; -  
 DR TIGR: VC0346; -  
 DR InterPro: IPR002627; IPT.  
 DR Pfam: PF01715; IPT; 1.  
 DR ProDom: PD004674; IPT; 1.

DR TIGRNAME: TIGR00174; m1aa: 1.  
 KW Transferase; Nucleotidyltransferase; tRNA processing; ATP-binding;  
 KW Complete proteome.  
 FT NP\_BIND 13  
 SQ SEQUENCE 315 AA; 35515 MW; 0114D280868CD876 CRC64;  
 Alignment Scores:  
 Pred. No.: 2, 93e-18 Length: 315  
 Score: 334.00 Matches: 99  
 Percent Similarity: 49.86% Conservative: 76  
 Best Local Similarity: 28.21% Mismatches: 116  
 Query Match: 9.34% Indels: 60  
 DB: 1 Gaps: 14

US-09-513-151-3 (1-2041) x M1AA\_VIBCH (1-315)  
 QY 65 CTGCAAGAGACCTCCTTGTAGTG---ATTCTGGGCGCACGGGCAATGCC 121  
 DB 1 MetThrGlnLysLeuProLeuAlaLeuPheLeuMetGlyProThrAlaSerGlyThr 20  
 QY 122 AGCGTGGCTTGACGCTAGAGCCAGCGGCTGGCGGTGACATGCTGACCGCTCCAG 181  
 DB 21 AspLeuAlaIleArgLeuArgGlnLysTyrProValGlnIleIleSerValAspSerAla 40  
 QY 182 CAGGTCTATGAGGCGCTAGACATCATCACCAACAAAGTTTCTGCCAAGAGCAGAAATC 241  
 DB 41 LeuIleTyrArgGlyMetAspIleGlyThrAlaLysProAspAlaGlnGluLeuAlaLeu 60  
 QY 242 TCCGCGCACCATCATGATGAGCTTGTGATCTCTGTGACCAATATACAGTGTGGAC 301  
 DB 61 AlaProHisArgLeuIleAspIleLeuAspProSerGlnAla---TyrSerAlaAlaAsp 79  
 QY 302 TTCAGAAATAGACCAACTGCTGATGAGATATATTTGCCGAGACAAATTCCTATT 361  
 DB 80 PheArgArgAspAlaLeuLysGlnMetAlaAspIleValAlaGlnGlyLysIleProLeu 99  
 QY 362 GTTGTGGAGGAGCAACATTTATTCATTTGATGATCTCTCTGCAAACTTTCTGCAATAC 421  
 DB 100 LeuValGlyGlyThrMetLeuTyrPheLysAlaLeu-----LeuGlnGlyLeu 115  
 QY 422 AAGCCCCAGAGATGGGCACTGAGAAATGATGACGGAAGTGGAGCTTGAAAAGAG 481  
 DB 116 SerProLeuPro---AlaAlaAspProValIleArgGlnGlnIleGluGlnAlaGln 134  
 QY 482 -----GAGGTGCTTGTTACTTCACAAACGCTTAAGCAGGTGACCCAGAAATGCTGCC 535  
 DB 135 LysLeuGlyTrpGlnAlaLeuHisAspGlnLeuGlnGlnIleAspProValSerAlaGln 154  
 QY 536 AAGCTGATCCACATGACAAAGCAAGAGTGGCAGAGCTTGCAATTTTGAAGAAACA 595  
 DB 155 ArgIleHisProAsnAspProGlnArgLeuSerArgAlaLeuGlnValTyrArgIleSer 174  
 QY 596 GGAATCTCATAGTAAGATTTCTCCATCGTCAACATAGCAGAAAGTGGTGGTCCCTT 655  
 DB 175 GlyLysThrLeuThrGlnLeu-----ThrGlnThrLysGlyAlaIle 189  
 QY 656 GAGGTCTCTGAGTCTCTACCCCTTGATCTTGGCTTCATGCTGACGAGCAGT 715  
 DB 190 -----ProTyrArgValLeuGlnPheAlaIle-----AlaProLysGlnTrpAlaGln 205  
 QY 716 CTAGATGAGCGTTGGATAGAGGGTGGATGACATGCTGCTGGCTTTGGAGGAA 775  
 DB 206 LeuHisArgArgIleGlnLeuArgPheGlnLysMetValGlnSerGlyPheGlnGlnGln 225  
 QY 776 CTAAAGATTTTCAACAGACGCTATATATCAGAAATGTTTCGAAATATGCGACGACTAT 835  
 DB 226 ValLysAlaLeuTyrAlaArg-----AspAsp 234  
 QY 836 CAACATGATATCTTCATCAATTT-----GGCTTCAGAAATTTTCACGAGTACTG 886  
 DB 235 LeuHisProAspLeuProSerIleArgCysValGlyTyrArgGlnMetTrpTyrLeu 254  
 QY 887 ATCACTAGAGGAAATGCACACTGGAG-----ACTAGTAAAC 922

```
Db 255 -----AspGlyHisGlyThrLeuAspGluAlaIleTyrArgGlyIleCysAlaThrArg 272
QY 923 CAGCTTCTAAGAAAGAGACCTGCTCCATTGTCCTATGCGCTTAGAGTACT 982
Db 273 GluLeuAlaLysArg-----GlnIleThr 280
QY 983 GATGCTCTGGAAGTGGAGAGACTCTGT-----CTTGAACCTGCTCTT 1024
Db 281 TripleuArgSerTrpAspAspLeuThrTripleuAspSerGluAsnValAspGluAlaVal 300
QY 1025 GAATCGTCGCAAGTTTCATCCAGGCCACAG 1057
Db 301 GluThrLeuSerAsnAlaIleAlaSerAsnGlu 311
```

Search completed: April 21, 2003, 18:49:17  
Job time : 32.423 secs